Win–Win Denial:
The Psychological Underpinnings of Zero-Sum Thinking

Samuel G. B. Johnson [1,2,3]
Jiewen Zhang [4]
Frank C. Keil [5]

[1] University of Warwick, Department of Psychology
[2] University College London, Centre for the Study of Decision-Making Uncertainty
[3] University of Bath, School of Management
[4] University of California – Davis, Department of Psychology
[5] Yale University, Department of Psychology

Corresponding Author:	Sam Johnson
Email for Correspondence:	sgbjohnson@gmail.com
Word Count:	16,320 (main text and acknowledgements)
Last Revised:	30 March 2021 (final version for copyediting)

In press at JEP: General.

Acknowledgements: We thank Gordon Brown, Clayton Critcher, Aparna Labroo, Angela Lee, Robert Meyer, Danny Oppenheimer, Stijn van Osselaer, and Klaus Wertenbroch for useful comments on this research, and audiences at MIT and the University of Bath for their feedback. A subset of these studies was presented at the Cognitive Science Conference (Johnson, Zhang, & Keil, 2018); we thank the reviewers and attendees for their comments.

Abstract

A core proposition in economics is that voluntary exchanges benefit both parties. We show that people often deny the mutually beneficial nature of exchange, instead espousing the belief that one or both parties fail to benefit from the exchange. Across 4 studies (and 8 further studies in the Supplementary Materials), participants read about simple exchanges of goods and services, judging whether each party to the transaction was better off or worse off afterwards. These studies revealed that *win–win denial* is pervasive, with buyers consistently seen as less likely to benefit from transactions than sellers. Several potential psychological mechanisms underlying win–win denial are considered, with the most important influences being *mercantilist theories of value* (confusing wealth for money) and *theory of mind limits* (failing to observe that people do not arbitrarily enter exchanges). We argue that these results have widespread implications for politics and society.

Keywords: Folk economics, zero-sum thinking, intuitive theories, theory of mind, decision-making.

Introduction

Economics is built on a seemingly commonsense insight: *There are gains from trade—voluntary transactions benefit both parties*. This has been recognized since at least the time of Adam Smith, who wrote in *The Wealth of Nations* (1999/1776):

> But man has almost constant occasion for the help of his brethren, and it is in vain for him to expect it from their benevolence only. He will be far more likely to prevail if he can interest their self-love in his favour, and show them that it is for their own advantage to do for him what he requires of them. *Whoever offers to another a bargain of any kind, proposes to do this. Give me that which I want, and you shall have this which you want, is the meaning of every such offer*; and it is in this manner that we obtain from one another the far greater part of those good offices which we stand in need of (p. 118–119; emphasis added).

Buyers do not buy something they value less than its price, and sellers do not sell something they value more. Within this range between the buyer's maximum and seller's minimum price, both parties benefit from the transaction. Arguably, opportunities for mutual gain from specialization and trade are the main reason why humans can accomplish more in groups than individually; the extent of these opportunities explains why economic development has proceeded more rapidly in some times and places than in others (Acemoglu & Robinson, 2012; Smith, 1999/1776).

The principle that voluntary transactions are win–win is a cornerstone of neoclassical economics (Mas-Colell, Whinston, & Green, 1995). But even behavioral economists—who have shown that people sometimes make suboptimal choices (Kahneman, 2002; Simon, 1955; Thaler, 1980 among many others)—agree that people make reasonable choices most of the time. For example, Thaler (1980) reassures us that "if a problem is sufficiently simple the normative [utility-maximization] theory will be acceptable" (pg. 39), citing demonstrations that key theorems of classical economics hold among irrational consumers (Becker, 1962) and even among rats (Kagel et al., 1975). Thaler concludes that heuristics lead to consumers to make "occasional" mistakes while they "are doing the best they can" (pg. 59). Indeed, a leading textbook in consumer behavior—hardly a field obsessed with praising human rationality—notes that "the idea that trade is always good is actually fairly obvious: if both parties were not better off, one or other would not be prepared to make the trade" (Blythe, 2013). Although people are not fully rational, they are not fools either: Absent coercion or deception, we rarely cede something we value highly for something we value less.

This principle implies that international trade is positive-sum—a conclusion endorsed by economists across the political spectrum, from Paul Krugman (1996) to Milton Friedman (1962). For example, when a panel of 44 ideologically diverse economists were asked to rate their agreement with the statement, "Freer trade improves productive efficiency and offers consumers better choices, and in the long run these gains are much larger than any effects on employment" 95% agreed or strongly agreed, and none disagreed (IGM Forum, 2012).

Yet many people seem to fallaciously deny the mutually beneficial nature of exchange (Caplan, 2002), and these attitudes appear to permeate public policy. Politicians who promote populist, anti-trade policies enjoyed enormous popularity on both left and right leading up to the 2016 U.S. presidential election. In 2018, the United States initiated trade wars with its two largest trading partners—China and Canada—and although the negative results were obvious ex ante to most economists, they seemed to surprise many laypeople. The fallacy of *win–win denial* may be ubiquitous.

Consistent with the possibility that win–win transactions are unintuitive, these intellectual battles have been fought before. Adam Smith was writing primarily in opposition to the mercantilist economic philosophy of his contemporary England—a now-debunked theory that conceptualized wealth as the accumulation of gold, rather than goods and services, and therefore emphasized the importance of exports. (Sound familiar?) Even though economists have been long convinced by Smith's arguments, battles against mercantilism and trade-protectionism must be fought anew each generation, as Ricardo (2004/1817), Bastiat (2007/1845), Marshall (1949/1879), Friedman (1962), and Krugman (1996) have done in turn. This need to re-learn basic economics anew each generation encourages the hypothesis that zero-sum thinking is psychologically natural—a hypothesis endorsed explicitly by economists including Bastiat (2007/1845) and Sowell (2008).

The denial of transactions as win–win fits can explain zero-sum thinking—the belief that one party's gain is another party's loss. Zero-sum thinking is usually mistaken in economics precisely because individual trades do not make individual parties worse-off. Yet it appears to be endemic in people's thinking about economic matters. Laypeople tend to believe that more profitable companies are less socially responsible (Bhattacharjee, Dana, & Baron, 2017), when the true correlation is just the opposite. Negotiators often perceive themselves as carving up a "fixed pie," decreasing the chances of a successful outcome (Bazerman & Neale, 1983; De Dreu, Koole, & Steinel, 2000). People believe that the government cannot benefit one group without harming another (Bazerman, Baron, & Shonk, 2001) and are particularly inclined to think in zero-sum ways about international trade (Baron & Kemp, 2004; Johnson, Zhang, & Keil, 2019) and immigration (Esses, Dovidio, Jackson, & Armstrong, 2001; Louis, Esses, & LaLonde, 2013). But zero-sum thinking also seems to be psychologically natural, occurring across many countries (Rózycka-Tran, Boski, & Wojciszke, 2015) and political orientations, though manifesting differently among liberals and conservatives (Davidai & Ongis, 2019). Zero-sum thinking has been noted in numerous settings (albeit not always fallaciously), including students' thinking about grades (Meegan, 2010), reasoners thinking about evidence (Pilditch, Fenton, & Lagnado, 2018), consumers' thinking about product features (Chernev, 2007; Newman, Gorlin, & Dhar, 2014), and even couples' thinking about love (Burleigh, Rubel, & Meegan, 2017).

Here, we examine how the denial of win–win transactions contributes to zero-sum thinking about simple transactions between individual buyers and sellers. We use this setting for three reasons. First, win–win denial is most clearly a normative error in simple economic transactions—like buying olive oil, hiring a hair stylist, or swapping kitchen supplies with a neighbor—compared to other, more complicated situations where there is more potential for behavioral biases to prevail or fraud to occur. (In the General Discussion, we consider whether, despite our best efforts, some of these transactions might be correctly perceived as zero-sum.) Second, these transactions are highly familiar: Given that consumers themselves have bountiful experience engaging in win–win trades, denial of mutual benefit here would be a particularly powerful demonstration of our hypotheses. Third, such simple transactions are nonetheless a microcosm of more complex transactions, such as international trade, that are at the heart of government policy-making.

Mechanisms

Our studies test when, and to what extent, people believe that one or both parties did not benefit from an economic exchange. We considered four possible contributing mechanisms.

Evolutionary Mismatch

First, some have argued that humans are evolutionarily adapted for like-kind exchanges such as barter (Boyer & Petersen, 2018; Cosmides & Tooby, 1992; Fiske, 1992; Pinker, 2003; see also Trivers, 1971). For example, people can adeptly solve otherwise unintuitive abstract reasoning problems such as the Wason card task when they are framed in terms of social exchange (Cosmides, 1989). It makes evolutionary sense that we should have mental modules devoted to regulating social exchange, which must include some mechanism for valuing potential exchange opportunities and ensuring that one benefits from each exchange. For instance, in Fiske's (1992) typology, Equality Matching relationships occur between trading partners when the exchange is "one-for-one" in the sense that similar items are reciprocally exchanged over time, as in gift exchange or carpooling.

However, it is not clear that these adaptations would be well-suited to contemplating modern, currency-mediated exchanges. Whereas Fiske's (1992) Equality Matching relationships appear to be near-universal across cultures, Market Pricing relationships—in which people trade disparate goods or services based on the principle of proportionality, typically mediated by money—seem to have become ubiquitous only after the Industrial Revolution (Polanyi, 1944). Money itself, of course, did not exist in our evolutionary ancestors' environment, nor did the requisite mathematics for market pricing. At best, we would expect intuitions to be worse for thinking about monetary exchanges rather than like-kind barters, where we should be adept at recognizing win–win trades. Conceivably, people might even undervalue money since it lacks intrinsic value, believing that buyers are better-off after transactions (since they gain valuable goods but give up intrinsically worthless paper), whereas sellers are worse-off.

Mercantilist Theories

Second, the error could lie in how people conceptualize value. Economists since Smith (1999/1776) have labored, with limited success, against mercantilist theories of wealth and trade. Such theories equate wealth with money, neglecting the insight that money is valuable only because it can be used to purchase valuable things. For instance, the United States is not made "poorer" by purchasing goods from China, as though imports are wealth-decreasing and exports are wealth-increasing. If China accepted dollars in exchange for products, and imported nothing in return—for example, because they burned the dollars—then Americans would be *wealthier*, not poorer, because they would have received something but would not have given up anything. Dollars are pieces of paper (or bits of information) with no intrinsic value in themselves. We are better off to the extent we can command resources, which may or may not be better facilitated by possessing currency or by exchanging it for useful commodities (e.g., see Hume, 1985/1752 on the balance of trade).

But there is reason to think that mercantilist thinking is widespread. Analogously, the medieval theory of motion (which posits curvilinear momentum for objects) was discredited by Newton and is now universally disclaimed among physicists. Yet this idea remains intuitive to laypeople and may even be built into our perceptual systems (Hubbard, 1996). Similarly, despite Smith's debunking of

mercantilism in the 18th century, it maintains an intuitive psychological appeal. In the General Discussion we speculate about reasons why mercantilism might be intuitive, but for now simply note that surveyed attitudes toward economic policies (Caplan, 2002) seem consistent with the hypothesis that laypeople, but not economists, rely on a mercantilist theory of value in their policy preferences.

In the context of our studies, mercantilism is equivalent to over-valuing money. If most people are intuitive mercantilists, then they would perceive sellers as better off than buyers, because the seller is always the party who gains currency, even if they had to give up something valuable in order to gain the currency. They would also tend to think that neither party benefitted from a barter, since neither party gained currency.

These predictions diverge strikingly from those made by neoclassical economic theory, particularly the prediction that sellers will be seen as winning from transactions whereas buyers will often be seen as losing. According to standard economic theory, welfare losses from voluntary transactions are not possible (Mas-Colell et al., 1995). And although models typically assume that it is theoretically *possible* for transactions not to benefit one or both parties, the party more typically getting the short end of these transactions is actually the *seller* in the conventional theory. Why? Because under perfect competition, sellers must price at marginal cost, experiencing zero economic profits (i.e., zero profits above taxes and the cost of capital; note that this does allow "accounting profits" that account for the alternative uses of the business's capital). Under these circumstances it is plausible to claim that the seller did not experience a welfare gain from a trade—although of course absolutely perfect competition is rare in the real world. The point is simply that buyers' *losses* from voluntary transactions are an emphatic departure from economic theory in the absence of force or fraud (an issue we consider in the General Discussion).

Theory of Mind

Third, the error could lie in cognitive processing—specifically, the failure to consider the mental states of the transaction parties. People often act as though their own subjective experiences reflect a reality that is equally accessible to others—a form of perspective-taking error known as naïve realism (Ross & Ward, 1997; see also Ross, Greene, & House, 1977 on the false consensus effect). Yet, one way to see that voluntary transactions are win–win is precisely to consider the preferences of the *buyer and seller*, rather than one's self. *You* may not think that a ticket to the Broadway musical *Hamilton* would benefit you at its $375 price tag, but when Russ buys a ticket at this price, he clearly values the ticket more than $375. Thinking about this from Russ's perspective makes this obvious, but naïvely believing the ticket has an intrinsic value less than $375 clouds the issue.

Thus, if theory of mind limits are implicated in win–win denial, experimental manipulations that encourage people to think about the transactors' mental states should make people more likely to view exchanges as mutually beneficial. People should be more prone to view transactions as win–win when the parties' reasons or preferences are made more explicit or when they are encouraged to take the perspective of the parties.

Heuristic Substitution

Fourth, the error could lie in how people approach the question. Even though economic theory and commonsense arguments do tell us the *direction* of welfare gains from trade, estimating the magnitude of these gains is very challenging. People might thus substitute this difficult question ("Is John better off after the exchange compared to before?") for a simpler question that we often consider as consumers ("Did John get a good deal, compared to other deals I know about?")—what we term the *bargain heuristic*. This attribute substitution process is thought to be fundamental to much heuristic judgment and decision-making (Kahneman & Frederick, 2002) by reducing the effort required to make any particular judgment even as errors can be introduced (Shah & Oppenheimer, 2008).

The reason that the bargain heuristic can lead to errors is that a purchase may very well be welfare-improving, even though welfare is not improved *as much* as it would be if the buyer got a better bargain. To tweak the famous example (Thaler, 1985), if you are staying at a remote resort hotel, the resort may well be able to use its local market power to charge you the exorbitant price of $10 for a beer. Yet if you consummate this exchange, it is precisely because you *still* value the beer more than $10 despite the fact that this is a poor bargain. The bargain heuristic will instead deliver the incorrect verdict that you were made worse-off by this trade. You are worse-off only compared to a counterfactual world in which the market is more competitive. You are *not* worse-off after the transaction than you were before.

Hence, if the bargain heuristic underlies win–win denial, then sellers should be seen as better-off and buyers often as worse-off when the price is high (a bad bargain). In contrast, when the price is low (a good bargain), buyers should be seen as better-off and, perhaps, sellers as worse-off, since in such cases the *seller* is getting the raw end of the bargain. This prediction differs from mercantilism, in that it does not place any special emphasis on money relative to real goods and services and outlines a case where sellers might be seen as worse-off (which should rarely occur on a mercantilist account, since sellers are gaining currency). The bargain heuristic account is related to the theory of mind account, since a perspective-taking error is a *prerequisite* to using the bargain heuristic. Thinking about the transaction from the transactors' perspectives makes it clear why the quality of the bargain is irrelevant to at least the *direction* of the welfare change—even if the buyer got a "bad" deal, it must not be so bad that she did not feel that she benefitted on balance.

Overview of Experiments

Four studies tested win–win denial and its moderators. The general method of these experiments was to ask participants about ordinary exchanges of goods or services—for example, Sally purchasing a shirt from Tony's store, Eric purchasing a haircut from Paul's barber shop, or Mark trading his soy sauce for Fred's vinegar. For each transaction, participants were asked whether or not each party was better off after the transaction. From the standpoint of neoclassical economics, *all* parties were better off after *all* exchanges, since people do not voluntarily enter into transactions at a loss, and we sought to avoid conditions under which behavioral amendments to economics would be likely to produce major exceptions. Nonetheless, if people engage in win–win denial, we would expect to see a widespread belief that some parties to these exchanges do not benefit.

The particular pattern of non-benefit can help to test the potential mechanisms for win–win denial. If mercantilism is the culprit, we would expect to see buyers (but not sellers) perceived as worse-off

and barters as pointlessly failing to benefit either party. On the other hand, the evolutionary mismatch account suggests that people may be better at recognizing positive-sum transactions among like-kind barters rather than monetary transactions, where people might even believe that sellers are made worse-off since they give up valuable goods in exchange for intrinsically valueless currency. These hypotheses were tested in Study 1.

Study 2 tested a further implication of mercantilism—that exchanges described in terms of time (labor) rather than money would be seen as more beneficial. Study 3 tested the theory of mind account by attempting to induce participants to take the perspective of the buyer by giving reasons for the buyer's purchase. Finally, Study 4 varied the prices of monetary exchanges to test heuristic substitution account, since very inexpensive products should then be seen as benefitting the consumers at the expense of the seller.

In the Supplementary Materials, we report several additional replication studies (Part B), including studies that varied the framing of the transactions or wording of the dependent variable (Studies S1, S4, and S5) and between-subjects replications of key results (Studies S2 and S3). We also pool data across studies to test individual differences in win–win denial (Part C), particularly educational and political predictors.

Study 1

Our first study tested win–win denial for various goods and services. We asked participants to read about simple, everyday transactions, including monetary purchases of goods (e.g., olive oil, a car), monetary purchases of services (e.g., a haircut, a plumber), and barters of goods (e.g., a McDonald's sandwich for a Burger King sandwich, or soy sauce for vinegar). Participants then rated the welfare of the buyer and seller (or traders, in the case of barter), relative to before the transaction.

This experiment probes two sets of questions. First, how *often* do people deny that transactions are win–win? If people understand the underlying principles of economics, they should indicate that both buyer and seller are better off after most or all transactions, because the transactions are voluntary. On the other hand, if people deny the win–win nature of trade, then they may often believe that either the buyer or seller failed to be bettered by the transaction, or even was worse off after the transaction.

Second, *what pattern* of perceived gains and losses do people perceive? The mercantilist theory of value—on which money has value over-and-above the goods and services it can purchase—suggests that benefits should "follow the money." That is, sellers should be seen as gaining more often than buyers (since sellers are gaining money and giving up a good or service, whereas the converse is true for buyers). Further, for barters, the traders would be likely seen as neither better *nor* worse off than before, because no money changes hands.

Methods

Across all studies reported in this article, we recruited U.S. participants from Amazon Mechanical Turk because they tend to be more diverse in age and education level compared to traditional undergraduate samples (albeit more politically liberal than the general U.S. public). Sample sizes were

set a priori and ranged from 100 participants (achieving 90% power for within-subjects effect sizes $d > 0.33$, using G*Power; Faul et al. 2007) to 200 participants (90% power for $d > 0.23$).

For Study 1, we recruited 100 participants ($M_{age} = 36.2$; 62% female, 43% college educated [i.e., having at least a 4-year college degree]) from Mechanical Turk; 14 were excluded from analysis because they incorrectly answered more than 33% of a set of comprehension questions (see below).

Participants read about a series of 12 transactions, and were instructed that "for each transaction, you will be asked whether each participant is better off, worse off, or the same, relative to how they were before the transaction." The transactions were divided into three types—monetary purchases of goods, monetary purchases of services, and barters of goods. Four items of each type were used, and the 12 items were presented in a random order.

For the monetary purchases of goods, participants read about transactions, such as "Sally goes to Tony's clothing store. She pays Tony $30 for a shirt." Other items included purchases of olive oil, a car, and a chocolate bar. Participants were then asked to rate the welfare-change of the buyer and seller—that is, how each party's welfare compares after versus before the transaction (e.g., "How well off do you think Sally now is?" and "How well off do you think Tony now is?") on a scale anchored at –5 ("Worse than before"), 0 ("Same as before"), and 5 ("Better than before"). Buyer and seller welfare-change were rated in a random order for each item. Monetary purchases of services were similar, except the transactions involved services rather than goods (e.g., "Eric goes to Paul's barber shop. Eric pays Paul $15 for a haircut."). Other items included purchases of massage, dog grooming, and plumbing services. For the barters of goods, participants read about two individuals exchanging goods, such as "Vivian goes to her colleague Tommy's office. She trades her Burger King hamburger for Tommy's McDonald's hamburger" or "Mark goes to his neighbor Fred. Mark trades his bottle of soy sauce for Fred's bottle of vinegar." Other items included trading a pencil for an eraser and swapping two designer bags. The welfare-change of each party was rated on the same scale as for the monetary transactions. The full text of all items is given in Part A of the Supplementary Materials.

After the main task, participants were asked a series of 12 check questions to monitor attention, each asking whether a particular item (e.g., "haircut" or "shirt") was mentioned on the previous pages. Participants incorrectly answering more than 33% of these questions were excluded from analysis.

At the end of the study, participants were asked to report demographic information. This included educational information, including education level, college major (if applicable), economics knowledge ("Please rate your knowledge of economics" on a scale from 0 ["Not knowledgeable"] to 10 ["Very knowledgeable"]), and number of economics courses. Participants were also asked about their political beliefs, including party affiliation, values on social issues, and values on economic issues, each on continuous 0 (Democrat/liberal) to 10 (Republican/conservative) scales. Finally, participants reported standard demographic information such as gender, age, and income. Part C of the Supplementary Materials reports descriptive statistics on these measures and their relationship with win–win denial.

Results

Win–win denial was endemic among our sample. Nearly all participants (94.2%) indicated that at least one party was made worse-off in at least one of the transactions. Since win–win denial is a logical prerequisite of zero-sum thinking—the belief that any gains from one party are compensated by losses

by the other party—it is not surprising that the vast majority of participants (88.4%) also believed that at least one of the transactions was zero- or negative-sum.

Figure 1 plots the proportion of times that buyers, sellers, and traders were deemed to have gained (the white area), lost (the black area), or experienced neither gain nor loss (the grey area) from each type of transaction. Clearly, people are not neoclassical economists who would color this whole chart white. But nor are they responding at chance. The patterns of win, loss, and no-benefit across groups are instructive for teasing apart the different theoretical explanations we proposed for win–win denial.

First, consider buyers versus sellers. Whereas basic economics says that both buyers and sellers benefit from transactions, people thought that buyers were much more likely to be made worse-off by their transactions than were sellers. Whereas very few sellers [$M = 0.49$, $SD = 0.86$ out of 8] were thought to be made worse-off by the trade, five times that many buyers were [$M = 2.53$, $SD = 2.47$ out of 8; $t(85) = 7.55$, $p < .001$, $d = 1.11$]. The same pattern was evident when looking at the mean welfare-change score rather than the trichotomized proportions [$M = 2.08$, $SD = 1.44$ vs. $M = 0.69$, $SD = 1.48$ out of 8; $t(85) = 8.75$, $p < .001$, $d = 0.95$]. This result is consistent with mercantilism, which predicts that value should be seen to flow with (over-valued) money.

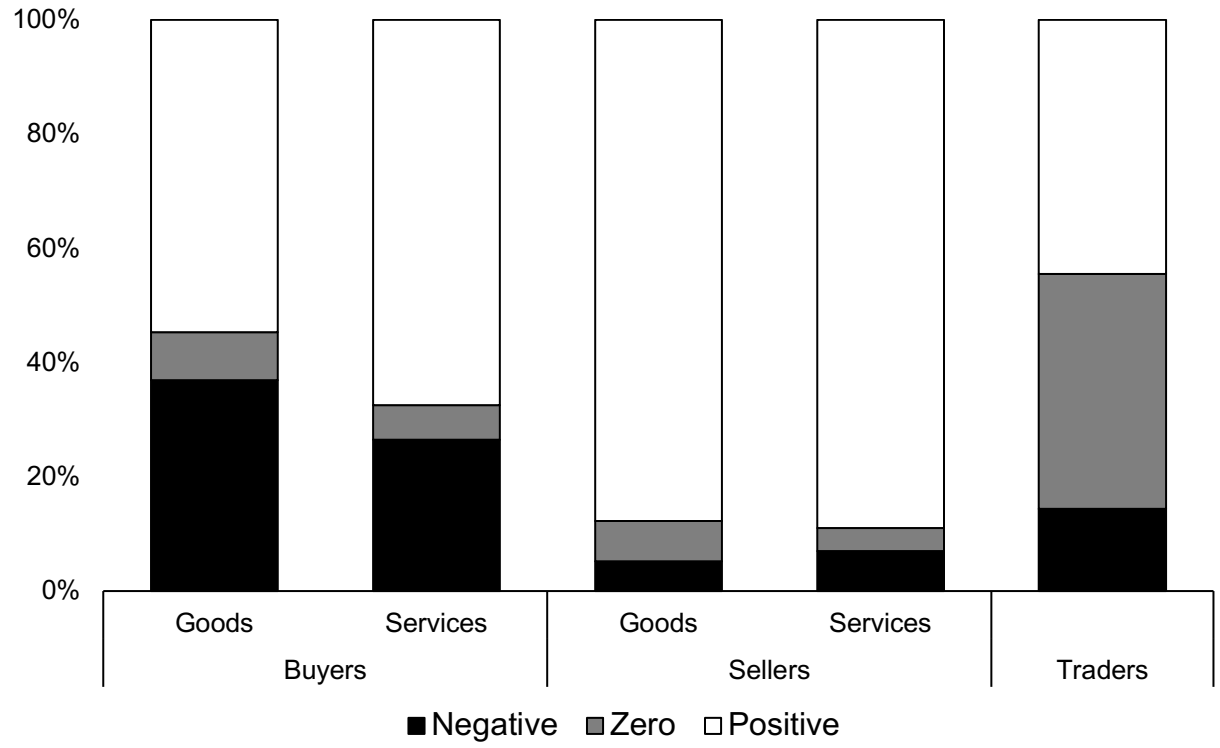

Figure 1. Proportion of transactions perceived as having negative, zero, or positive impact on buyers, sellers, and traders in Study 1.

Second, consider monetary transactions versus barters. Economics says that both traders in a barter benefit, since they otherwise would not have decided to make the trade. In contrast, participants were even more likely to perceive traders as failing to benefit compared to buyers. This occurred because the proportion of no-benefit transactions for traders [$M = 3.29$, $SD = 2.98$ out of 8] was perceived as far higher for traders than for either buyers [$M = 0.58$, $SD = 1.35$; $t(85) = 8.59$, $p < .001$,

d = 1.17] or sellers [M = 0.44, SD = 1.14 out of 8; $t(85) = 8.76$, $p < .001$, d = 1.26]. This led to a less positive welfare-change score for traders than for monetary transactions (averaging across buyers and sellers) [M = 0.51, SD = 0.93 vs. M = 1.38, SD = 1.26; $t(85) = 7.15$, $p < .001$, d = 0.79]. Once again, this is consistent with mercantilism, which predicts that transaction parties should be seen as gaining more when they are gaining currency rather than goods: On this mechanism, barters should often be seen as involving no-change in welfare to either party. This contrasts with the predictions of the evolutionary mismatch mechanism, which posits that we are adapted to reason about barters. If anything, people should thus be more prone see barters as win–win, since this view maintains that we are adapted for like-kind exchange.

The more pronounced win–win denial for barters translated directly into zero-sum thinking. For the monetary transactions, participants thought on average that 0.90 (SD = 1.11) out of the 4 purchases of goods and 0.62 (SD = 1.04) out of the 4 purchases of services to be zero- or negative-sum, in the sense that the sum of both parties' welfare-change scores was either zero or negative. These beliefs, while not overwhelming in magnitude, were widespread among participants: Most participants (60.5%) claimed that at least one of the monetary transactions was zero- or negative-sum. But zero-sum thinking for barters was much more common: More than half of the barters [M = 2.13, SD = 1.50 out of 4] were deemed zero- or negative-sum, far more than for either monetary purchases of goods [$t(85) = 7.46$, $p < .001$, d = 0.93] or services [$t(85) = 9.20$, $p < .001$, d = 1.17]. The great majority of participants (79.1%) believed that at least one of the barters was zero- or negative-sum.

An alternative way to look at zero-sum thinking considers that a '+3' from one party and '–2' for the other does not necessarily reflect positive-sum thinking, as the preceding analysis assumes, as these welfare change scores are not necessarily commensurable. Thus, despite the considerable magnitude of zero-sum thinking revealed there, this may actually understate its prevalence. Table 1 reports the proportion of the 8 monetary transactions and 4 barters in which each party was seen as gaining, losing, or experiencing no change in welfare. Part A in the Supplementary Materials reports these results broken down by item, revealing that the pattern generalizes across items.

		Sellers		
		Negative	**Zero**	**Positive**
Buyers	**Negative**	1.45%	0.44%	29.80%
	Zero	1.02%	3.05%	3.20%
	Positive	3.63%	2.03%	55.38%
		Traders		
		Negative	**Zero**	**Positive**
Traders	**Negative**	2.33%	2.03%	22.10%
	Zero		36.92%	6.40%
	Positive			30.23%

Table 1. Transaction types in Study 1.

Note. Entries are the proportion of transactions, analyzed pairwise, in which each party was seen as benefitting ("positive"), losing ("negative"), or experiencing no change in welfare ("zero").

For monetary transactions, participants agreed that about half of the transactions (55.4%) were win–win. Of the transactions denied to be win–win, a large majority were seen as win–lose (29.8% of the total) with sellers benefitting at the expense of buyers, whereas the converse (lose–win, with sellers losing but buyers benefitting) were rare (3.6%). This is consistent with the notion that mercantilism translates into rampant zero-sum thinking, but at odds with mainstream economic theory, according to which buyers are, if anything, *more* likely to gain from trade than sellers from voluntary transactions in competitive markets. Other combinations—such as negative-sum thinking with both parties losing or thinking that either the buyer or seller experienced no change in welfare—were rare, together accounting for about 10% of the transactions.

For barters, participants denied that *most* of the transactions were win–win, with only 30.2% seen as win–win. Of the transactions not seen as win–win, about half were seen as benefitting neither party (36.9% of the total), and a substantial minority were seen as benefitting one trader but not the other (22.1%). Other combinations were again rare, together accounting for about 10% of the transactions. This again accords with the prediction of mercantilism that people should often deny any change in welfare when money does not change hands.

Discussion

Study 1 revealed three key results. First, a large majority of participants espoused win–win denial and a zero-sum mentality at least some of the time. This is consistent with our suggestion that these are natural psychological tendencies, supporting claims (e.g., Caplan, 2007; Rubin, 2003) that zero-sum thinking accounts for at least a part of the difference between economists' and laypeople's reasoning about markets.

Second, within monetary transactions, sellers were almost always seen as net beneficiaries, whereas buyers were often seen as net losers. This way of thinking seems bizarre, because it seems to imply that consumers believe themselves to be acting irrationally when making purchases—if purchases are often net losses, why do consumers make them? Yet, this pattern is consistent with mercantilist theories that equate money and wealth: Apparently Adam Smith (1999/1776) did not slay this demon.

Third, win–win denial was even more prominent for barters than for monetary transactions: Traders were very often seen as neither gaining nor losing from their barters. This result further undercuts the idea that win–win denial is due to evolved instincts about exchange: To the extent that our evolutionary ancestors engaged in explicit trade of goods (e.g., in reciprocal altruism; Trivers, 1971), these would have taken the forms of barters rather than monetary transactions. On this hypothesis, people should, if anything, be *more* prone to deny that monetary transcations are win–win—the opposite of our finding. Of course, this evidence does not undermine the broader claim that humans are adapted for barter exchanges but not for money (e.g., Pinker, 2003)—only the idea that this adaptive mismatch is a major driver of win–win denial.

One possible concern is that participants interpreted the phrases "better off" and "worse off" as referring specifically to *financial* well-being. We address this concern in three ways. First, in Study S1 in the Supplementary Materials, we repeated Study 1 but phrasing the dependent variable in terms of whether or not parties "benefited" from the transaction. This study revealed a smaller degree of zero-

sum thinking overall, but comparable asymmetries between buyers versus sellers and between monetary transactions versus barters. Second, in Study S4 in the Supplementary Materials, we add detailed instructions that emphasize that "better off" and "worse off" are not defined solely in terms of money but in terms of broader utility, as well as a series of check questions to ensure that participants understand this definition. The results of that study are very similar to Study 1. Finally, we will see that this interpretation of the wording would not lead to the differences we observe across subsequent experiments, since this phrasing was kept constant across the remaining experiments.

Could one argue that these results actually *contradict* the notion of win–win denial, since a great majority of sellers, modest majority of buyers, and nearly half of traders were seen as benefitting from the transaction? We think this is a tough case to make, because chance responding is not a relevant comparison: The normative theory says that all (or nearly all) of the transactions should be seen as mutually beneficial, thus *that* is the most appropriate comparison. Empirically, we observe that *nowhere near* all of the transactions were seen as win–win. Consider base rate neglect—people's inattention to prior probabilities—as a comparison. Whereas early studies seemingly found that people do not use base rates at all, more careful work later on revealed that people often do rely on them to some degree (see Koehler, 1996). Yet, it is still a mistake when people underweight base rates, even if they do not fail to consider them entirely. Similarly, it is a fallacy when people deny the mutually beneficial nature of trade even as they do not fail to recognize this entirely. These glasses can be plausibly be viewed as half-full (since people often use base rates to some degree and often acknowledge that trades are mutually beneficial) or half-empty (since people systematically underweight base rate information and systematically underappreciate the gains from trade). But these glasses are definitely not *full*.

Ultimately one's conclusion about our participants' reasonableness will depend on one's acceptance that the transactions in our studies should be properly viewed as win–win—that neoclassical economics is the appropriate normative standard for these cases. We think it is: Notwithstanding the existence of irrationality, most of our transactions were as routine and transparent as typical market transactions, and few would argue that behavioral bias afflicts all market activity (see Thaler, 1980). We flesh out this argument more fully in the General Discussion. Nonetheless, even if one believed that our participants' level of win–win denial was reasonable, one would still need to explain its underlying mechanism. The current results support mercantilism in that benefit was widely seen as flowing with currency.

Given the fairly high levels of win–win denial we see in the current studies, it is natural to ask whether this habit of thought is equally widespread among all participants, or if educational, ideological, or demographic differences moderate its magnitude. In Part B of the Supplementary Materials, we pool data from all studies (over 1000 participants) to test for individual differences with a great degree of statistical power. There were modest effects of economics knowledge, with more knowledgeable participants somewhat less prone toward win–win denial, but perhaps surprisingly, political ideology had little effect, nor did any demographic factor we tested. Thus, win–win denial does not seem to be tied to any particular political ideology and its levels are high even among those with substantial economics training. This speaks to its deep psychological roots.

Study 2

Study 1 provided several key pieces of evidence for mercantilist thinking as a mechanism for zero-sum thinking: Individuals were almost always considered better off if they gained currency in a transaction (sellers), often considered worse off if they gave up currency (buyers), and usually considered no better off when no currency changed hands (traders). A further prediction of mercantilism is that taking the focus off of currency should reduce the asymmetry between buyers and sellers. That is, even though money usually changes hands in transactions, money is merely an intermediary that allows us to trade multilaterally—receiving money (usually in exchange for labor to our employer) and paying money (in exchange for goods and services from merchants) with different counterparties. Otherwise, surgeons wishing to get a haircut would need to search for a barber in need of a bypass operation, and professors wishing to purchase groceries would need to search for grocers in need of social psychology lessons. Money is a "veil" that facilitates barters of goods and services across individuals and over time.

Though explainable intuitively, this idea may not occur spontaneously to people, which may in part explain the appeal of mercantilist thinking and zero-sum beliefs. It may nonetheless be possible to encourage this way of thinking by highlighting costs in terms of the labor required to make purchases, rather than in explicit monetary terms. We tested this prediction in Study 2 by describing transactions in terms of time (i.e., the number of hours worked to earn sufficient money to make a purchase or to earn the good being bartered) rather than the money or good itself. De-emphasizing that transactions are mediated by money should reduce the perception that buyers are made worse-off (since their money-loss is less salient) but also the perception that sellers are made better-off (since their money-gain is less salient).

Methods

We recruited 200 participants (M_{age} = 36.4, 47% female, 54% college educated) from Mechanical Turk; 33 were excluded from analysis based on the criteria used in Study 1.

The procedure was similar to Study 1, except each item appeared in either the *money-frame* or the *time-frame*. The money-frame version of each item was the same wording used in Study 1, whereas the time-frame version described the buyer's (or one trader's) side of each transaction in terms of labor rather than money. For example, the money-frame version of one item read "Sally goes to Tony's clothing store. She pays Tony $30 for a shirt," whereas the time-frame version read "Sally goes to Tony's clothing store. Sally worked for 1.5 hours to pay Tony for a shirt," and in a separate paragraph noted that, "Sally gets paid $20 per hour (after taxes) at her job." This is mathematically equivalent to the Study 1 version, since 1.5 hours of work at $20 per hour equals $30. This equivalency was maintained across all items.

A parallel modification was made for the barter items. For example, one Money-frame item read "Vivian goes to her colleague Tommy's office. She trades her Burger King hamburger for Tommy's McDonald's hamburger," whereas the time-frame version read "Vivian goes to her colleague Tommy's office. Vivian worked for 20 minutes to pay for a Burger King hamburger, which she trades to Tommy in exchange for Tommy's McDonald's hamburger." Vivian's hourly wage was not included, since the explicit monetary cost was not included in Study 1.

Two items from each category (monetary purchases of goods, monetary purchases of services, and barters) appeared in each of the two conditions (counterbalanced). The money- and time-frame items were blocked, appearing in separate halves of the experiment, in a counterbalanced order. Within each block, items were presented in a random order.

Results

Overall, framing the transactions in terms of money versus time led to shifts in the perception of who gained versus lost from the transactions. While zero-sum thinking remained endemic and buyers were still more often seen as worse-off compared to sellers, these tendencies were muted.

We look first at the money-frame, which replicated the basic results of Study 1. As shown in Figure 2, participants were much likelier to think that buyers compared to sellers had negative outcomes from their transactions [M = 1.33, SD = 1.38 vs. M = 0.21, SD = 0.51 out of 4; $t(166) = 9.65$, $p < .001$, d = 1.07]. Consequently, the average welfare score was much lower for buyers than for sellers [M = 0.67, SD = 1.53 vs. M = 1.92, SD = 1.31; $t(166) = 10.21$, $p < .001$, d = 0.88].

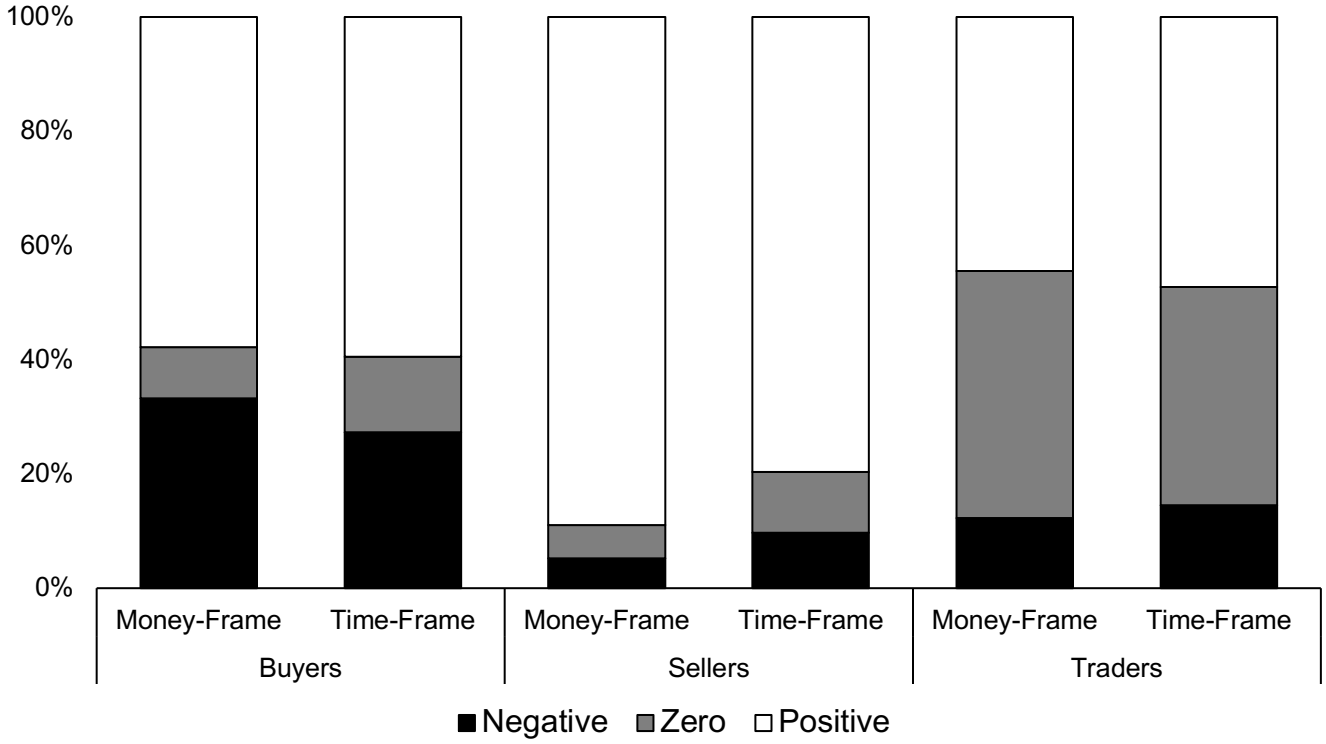

Figure 2. Proportion of transactions perceived as having negative, zero, or positive impact on buyers, sellers, and traders across the money-frame and time-frame conditions of Study 2.

Comparing trades versus monetary transactions, participants considered traders as much likelier to experience zero change in welfare [M = 1.73, SD = 1.66 out of 4] compared to buyers [M = 0.36, SD = 0.81 out of 4; $t(166) = 10.87$, $p < .001$, d = 1.05] or sellers [M = 0.23, SD = 0.70 out of 4; $t(166)$ = 11.17, $p < .001$, d = 1.18]. This led to less positive perceived welfare-change for barters than for monetary transactions (averaging across buyers and sellers) [M = 0.65, SD = 1.17 vs. M = 1.29, SD =

1.18; $t(166) = 6.87$, $p < .001$, $d = 0.55$]. Overall, the results for buyers, sellers, and traders in the money-frame replicate Study 1 not only qualitatively, but also reveal very similar means and effect sizes.

If mercantilism indeed explains the asymmetries between buyers and sellers, we would expect time-framing to attenuate the perceived advantages to sellers, since time-framed transactions do not emphasize the transfer of (over-valued) money. Indeed, the trends observed for the money-frame were less pronounced in the time-frame, as shown in Figure 2. Compared to the money-frame, buyers in the time-frame were seen less often as losing [$M = 1.33$, $SD = 1.38$ vs. $M = 1.09$, $SD = 1.31$ out of 4; $t(166) = 2.42$, $p = .017$, $d = 0.18$], and sellers less often as gaining [$M = 3.56$, $SD = 3.19$ vs. $M = 3.19$, $SD = 1.20$ out of 4; $t(166) = 4.26$, $p < .001$, $d = 0.36$]. Relative to the money-frame, the time-frame thus led to (marginally) higher average welfare scores for buyers [$M = 0.67$, $SD = 1.53$ vs. $M = 0.89$, $SD = 1.52$; $t(166) = 1.96$, $p = .052$, $d = 0.14$] and lower scores for sellers [$M = 1.92$, $SD = 1.31$ vs. $M = 1.40$, $SD = 1.29$; $t(166) = 5.12$, $p < .001$, $d = 0.40$]. Since money did not change hands, framing barters in terms of time-equivalent units should not influence judgments of welfare, and indeed this manipulation had little effect on perceptions of traders' welfare: Relative to the money-frame, trades were only marginally less likely to be seen as zero-gain [$M = 1.73$, $SD = 1.66$ vs. $M = 1.53$, $SD = 1.62$ out of 4; $t(166) = 1.66$, $p = .099$, $d = 0.12$] and the overall welfare-change score did not differ [$M = 0.65$, $SD = 1.17$ vs. $M = 0.66$, $SD = 1.14$; $t(166) = 0.18$, $p = .86$, $d = 0.01$].

Discussion

Overall, these results further buttress the mercantilism account. In addition to replicating the results of Study 1, Study 2 revealed that encouraging participants to think in terms of the time-cost rather than money-cost of transactions decreased win–win denial, with buyers about 18% less likely to be seen as losing. This is consistent with mercantilism, which predicts that any framing that de-emphasizes money would decrease the asymmetry between buyers' and sellers' perceived welfare.

A possible objection to these findings is that money in fact played no smaller role in these exchanges, but instead the transactions were merely framed in an unusual way. However, we would argue that this is precisely the point of this manipulation—to establish that drawing attention away from money and toward the more fundamental resource being exchanged (time) reduces the tendency toward win–win denial. A counterpoint may be that the framing manipulation is confounded in some way, and not only shifts attention but also introduces demand characteristics, changes inference patterns, or adds noise. For example, a reminder that one was willing to work in exchange for a good might increase the perceived strength of preferences for that good, or the unusual framing may have disrupted typical intuitions and added noise that makes the pattern less extreme.

Studies we report in the Supplementary Materials speak against these possibilities. Study S2A replicates the moderating effect of time-framing in a between-subjects setting, speaking against potential demand characteristics. Study S2B frames the transactions in terms of both time and money; this should equally remind participants that the person had to work in exchange for the good and is similarly an unusual framing. Nonetheless, these results look very similar to the money-frame. This suggests that the time-frame reduces win–win denial not because it emphasizes time, but because it *de-emphasizes* money, supporting the argument for mercantilism.

Nonetheless, mercantilism alone does not appear to *fully* explain zero-sum thinking. Even with an emphasis on time rather than money, win–win denial was rampant in the time-frame condition, with more than 25% of the buyers still thought to be worse-off after their transactions. Thus, Studies 3 and 4 examine other potential mechanisms.

Study 3

The key for why most transactions are win–win is that these transactions are *voluntary*, and people do not generally make purchases without believing it to benefit them. Economists do not need to be reminded of this, but laypeople might: Perspective-taking is effortful (Davis, Conklin, Smith, & Luce, 1996; Epley, Keysar, & Van Boven, 2004) and we erroneously impute our own beliefs and preferences to others (Ross & Ward, 1997; Ross et al., 1977). That is, absent a nudge to consider the buyers' and sellers' mental states, people may simply substitute their own preferences, which are likelier to be less favorable toward a purchase than the preferences of a person who is known to have made that purchase. Thus failure of perspective-taking could lead people to fail to realize that a transaction is unlikely to be conducted at a loss to either party.

Study 3 tested whether giving *reasons* for the buyers' decisions would attenuate win–win denial by making the voluntary nature of the transactions salient. Although we test other reasons in the Supplementary Materials (Study S3), Study 3 focuses on "empty reasons"—merely indicating that the buyer "wanted" the good or service being purchased (cf. Langer, Blank, & Chanowitz, 1978). These reasons are empty, from an economic perspective, because it is true of *every* voluntary purchase that the consumer wanted to purchase the product. Yet these reasons may not be psychologically empty, if they induce the participant to take the perspective of the buyer as a voluntary agent who would not choose to make a purchase at a loss.

Methods

We recruited 198 participants (M_{age} = 38.8; 51% female; 55% college educated) from Mechanical Turk; 23 were excluded from analysis based on the criterion used in previous studies.

The procedure was similar to Study 1, except each item appeared in either the *no-reason* or *empty-reason* condition. The no-reason version of each item used the same wording as Study 1, whereas the empty-reason version included a statement that the buyer or traders "wanted" to carry out the transaction. For example, Sally's and Eric's monetary transactions were explained as, "Sally made the purchase because she wanted the shirt" and "Eric made the purchase because he wanted the haircut." Mark's and Fred's barter was explained as "Mark made the trade because he wanted vinegar and Fred made the trade because he wanted soy sauce."

Analogously to the design of Study 2, two items from each category appeared in each of the two conditions (counterbalanced). The conditions were blocked and appeared in a counterbalanced order, with items presented in a random order within each block.

Results

As shown in Figure 3, the same basic asymmetries between buyers, sellers, and traders were seen in Study 3. But these asymmetries were significantly smaller when a reason (albeit empty) was given for the buyer's or traders' choice.

If theory of mind limits partly account for win–win denial, we would expect buyers to be seen as likelier to benefit when participants are encouraged to think about the buyers' reasons. Indeed, buyers were seen as less likely to lose from transactions in the empty-reason than in the no-reason condition [$M = 0.87$, $SD = 1.25$ vs. $M = 1.09$, $SD = 1.41$ out of 4; $t(174) = 2.43$, $p = .016$, $d = 0.16$], manifesting in a comparable trend in the welfare-change scores [$M = 1.32$, $SD = 1.57$ vs. $M = 0.86$, $SD = 1.64$; $t(174) = 4.64$, $p < .001$, $d = 0.29$]. Thus, an intervention as simple as mentioning that the buyers "wanted" the products they were purchasing significantly attenuated win–win denial for buyers.

Analogously, we would expect traders to be seen as likelier to benefit when their reasons are given. As shown in Figure 3, the proportion of "zero" responses was far smaller in the empty-reason than in the no-reason condition [$M = 1.54$, $SD = 1.60$ vs. $M = 0.98$, $SD = 1.46$ out of 4; $t(174) = 4.53$, $p < .001$, $d = 0.37$], resulting in much higher average welfare judgments for traders in the empty-reason than in the no-reason condition [$M = 1.84$, $SD = 1.80$ vs. $M = 0.81$, $SD = 1.20$; $t(174) = 8.32$, $p < .001$, $d = 0.67$]. This again supports theory of mind limits as a mechanism for win–win denial, with perspective-taking a partial remedy.

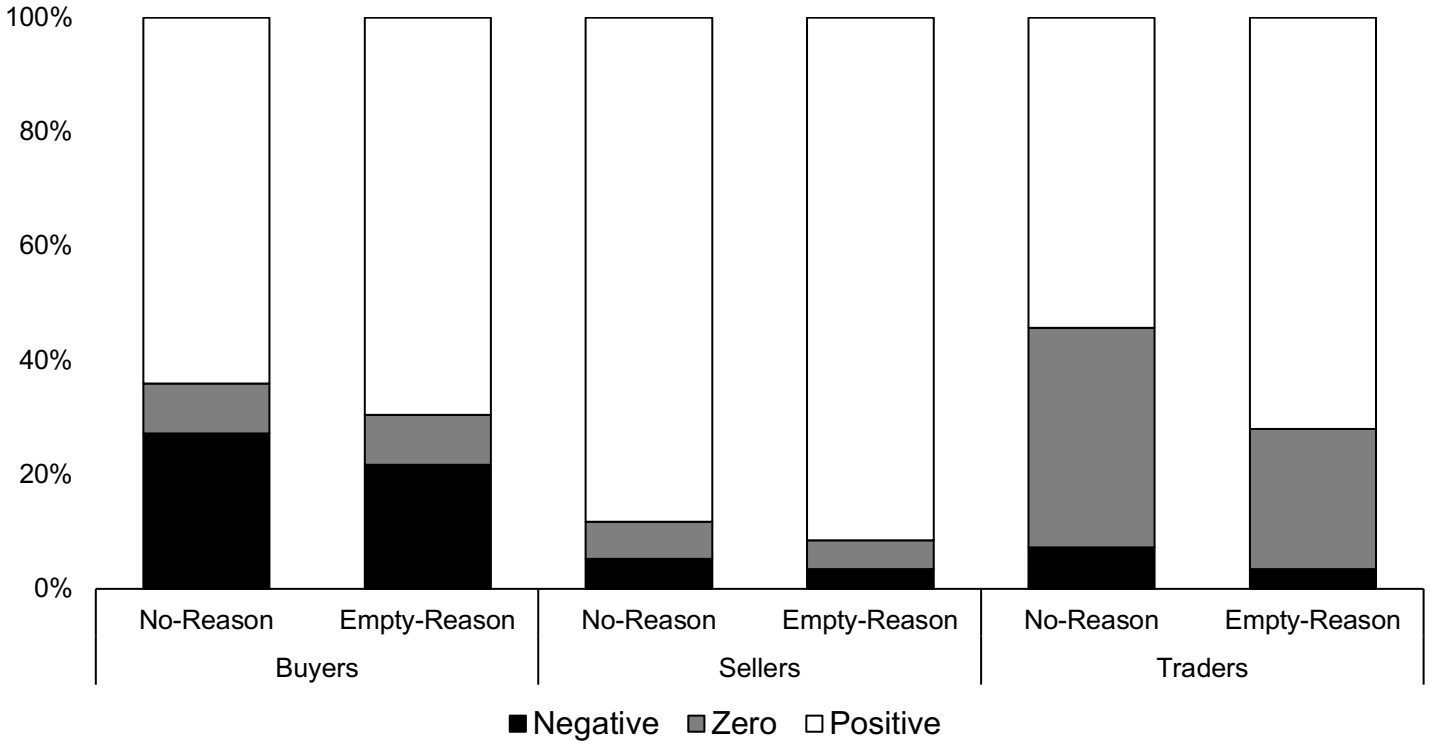

Figure 3. Proportion of transactions perceived as having negative, zero, or positive impact on buyers, sellers, and traders across no-reason and empty-reason conditions of Study 3.

We would not necessarily expect a comparable effect for sellers, for two reasons. First, the scenarios only mentioned the buyer's reasons. Second, the seller's reasons—profit-maximization—may be more obvious and therefore less in need of emphasis. Indeed, people automatically perceive

nefarious motives among sellers, ascribing to them powers of manipulation that are out of step with reality (Khon, Johnson, & Hang, 2020). Nonetheless, there was a trend for sellers to be seen as more likely to gain in the empty-reason than in the no-reason condition [M = 3.66, SD = 0.71 vs. M = 3.53 out of 4, SD = 0.83; $t(174) = 2.06$, $p = .041$, $d = 0.17$], albeit a trend that did not reach significance for average welfare ratings [M = 1.88, SD = 1.32 vs. M = 2.01, SD = 1.43; $t(174) = 1.52$, $p = .13$, $d = 0.10$]. One possible explanation is that contemplating the buyers' reasons for their purchases cued broader mentalizing that also led participants to consider the sellers' reasons as well.

Discussion

Study 3 found that cueing participants to consider the buyers' and traders' reasons for their transactions led them to recognize that those transactions were mutually beneficial: Buyers were 20% less likely to be seen as losing and barters were 36% less likely to be seen as inducing no benefits to either party. This implicates theory of mind limits as a partial explanation for win–win denial, as participants who treat their preferences as reflective of reality may not recognize that others' decisions are guided by their own reasons and preferences.

In the Supplementary Materials, we report several further studies (Studies S3A–B, S4, and S5A–B) that are consistent with Study 3. We replicated the reduction in win–win denial when the empty-reasons condition is tested in a between-subjects setting (Study S3A); indeed, the reduction was considerably larger there than in the within-subjects design of Study 3, underscoring the robustness and magnitude of the effect. A comparable effect was also seen when more explicit, preference-based reasons were given (e.g., Sally bought the shirt because she saw Taylor Swift wear it) (Study S3B).

Study S4 used a different manipulation to draw participants' attention to the perspective of the people undertaking the transaction—asking participants to indicate the preferences of each party (e.g., indicating that the buyer preferred the goods over the money exchanged while the seller had the opposite preferences). This manipulation reduces win–win denial compared to a control condition, with this effect stronger than that in Study 3. Even here, though, participants were more prone to claim that buyers lose from exchanges compared to sellers, claiming that buyers "lose" about one-fourth of the time even after explicitly stating that the buyers preferred the good over the money.

Study S5 tested two other manipulations which further implicated theory of mind. Study S5A asked participants to explicitly rate welfare from the buyers' and sellers' perspective (e.g., "How well off do you think Sally believes she now is?"). As predicted by the theory of mind account, this led to less win–win denial compared to Study 1, although this was less effective than giving reasons.

Study S5B, in contrast, asked participants to imagine *themselves* as the buyers and traders (e.g., "You go to Tony's clothing store…"). On the one hand, people have high opinions of their own abilities (e.g., Dunning, Meyerowitz, & Holzberg, 1989) and their decision-making faculties in particular (Johnson & Rips, 2014, 2015; Pronin & Kugler, 2010) and might therefore be more likely to think themselves beneficiaries from their transactions. But the theory of mind account arguably makes the opposite prediction. Participants know their own preferences, which might *indeed* value money more than the good they are buying. Thus, if theory of mind limits explain the high rates of win–win denial in our other studies, we would expect a similar result whether transactions are framed in first-person or third-person. This is in fact what Study S5B found.

Study 4

Assessing relative welfare before and after a transaction is not an easy problem to solve. We can say, normatively, what the *direction* of this change will be, because people do not engage in voluntary transactions unless they believe they will benefit. Yet, the precise gain in welfare is elusive even to economists because it depends on consumers' subjective valuation, which is difficult to measure. In other contexts where judgments are difficult, people often substitute the hard question for an easier question (Kahneman & Frederick, 2002). In this case, an easier question would be: *Did the consumer get a good deal?* If people use this heuristic, they may not compare the buyers' relative welfare *before and after the exchange*, but instead the quality of the buyer's bargain *relative to other possible bargains*. That is, people may use a *bargain heuristic*.

To test this possibility, Study 4 manipulated the price of the monetary transactions, varying it from 50% of the approximate market price, up to 150%. As the price decreases, the buyers do indeed benefit more, and the sellers less, relative to a higher price (in economists' language, this changes the extent to which the gains from trade contribute to the *consumer surplus* versus the *producer surplus*). However, this does not change the economists' conclusion that both buyers and sellers benefit from *any voluntary transaction*, regardless of the price: A high price merely indicates that the buyer and seller placed a relatively high valuation on the item. Thus, the bargain heuristic is one particular manifestation of a theory of mind error, in that thinking about the transaction from the buyers' and sellers' perspective makes plain that the quality of the bargain is not relevant to the *direction* of the welfare change. A failure in perspective-taking is a prerequisite to using this heuristic.

If people use the "bargain" heuristic, they should no longer think that the buyers lose from transactions at low prices, and may even believe that the sellers lose at these prices. Conversely, if people endorse zero-sum transactions primarily due to mercantilism, they would often believe that buyers lose from transactions even when they get a good deal. Finally, if these mechanisms combine to produce win–win denial, buyers should be seen as losing more often than sellers at all bargain levels, but this difference might be attenuated for particularly good bargains and exacerbated for particularly bad ones.

Methods

We recruited 150 participants (M_{age} = 38.6; 56% female, 55% college educated) from Mechanical Turk; 8 were excluded from analysis based on the criterion used in previous studies.

The procedure was the same as Study 1, except that ten items were used—five monetary purchases of goods and five monetary purchases of services (barter items were excluded because monetary price could not be manipulated). Some of the items differed from previous experiments to avoid inferences about price/quality trade-offs (e.g., a car purchase could not be used, because a discounted car would likely be assumed to be used).

Among the five monetary purchases of goods and the five monetary purchases of services, participants read about one good and one service at typical market price (e.g., "Eric goes to Paul's barber shop. Eric pays Paul $15 for a haircut"), as well as one each at 50%, 75%, 125%, and 150% of the market price (i.e., for the haircut, $8, $11, $19, or $23). Participants were not told what the market

price or discounts were—conditions differed only in the dollar amount of the transaction. The five price conditions were counterbalanced with the five different goods and five different services using Latin squares, so that each item was paired with one of the five price conditions for each participant, and items were paired equally often with price conditions across participants.

Results

Although participants were sensitive to price, believing that buyers were comparatively more likely to be the same or worse off when the prices were high, they continued to deny that buyers gained even when the price was 50% of the market price (Figure 4). This suggests that a "bargain" heuristic, while perhaps contributing to win–win denial, is not all there is to the story.

To test whether price influenced judgments of buyers' welfare, we calculated a linear contrast for each participant, comparing the number of times buyers were thought to have negative welfare-change across price levels. These contrasts were significantly positive [$M = 0.63$, $SD = 1.50$; $t(141) = 4.98$, $p < .001$, $d = 0.42$], indicating that people were likelier to think the buyers were harmed by transactions at higher price levels. The equivalent contrasts on the raw means were significantly negative [$M = -1.33$, $SD = 3.32$; $t(141) = -4.76$, $p < .001$, $d = 0.40$], reflecting lower perceived welfare at higher price levels. Note that the latter result alone is appropriate according to economic theory, since the gains from trade are indeed apportioned more to sellers when the price is high and more to buyers when the price is low. But the former result is not consistent with economic theory—at no price do buyers regularly lose from trade, assuming that buyers voluntarily transacted at that price.

The results for sellers were less robust. We computed linear contrasts, comparing the number of times sellers were thought to have positive welfare-change across price levels. These contrasts, while numerically positive, were not significantly different from zero [$M = 0.18$, $SD = 1.34$; $t(141) = 1.57$, $p = .12$, $d = 0.13$]. Thus, sellers were not seen as particularly likelier to benefit at higher price levels. Yet, this could be due to ceiling effects, since (as is obvious from Figure 4) sellers were nearly always seen as gaining from trades. This suspicion is confirmed by looking at the contrasts calculated on the means rather than number of positive welfare-change scores, which were significantly positive, to a similar degree as buyers' mean scores were negative [$M = 1.30$, $SD = 2.86$; $t(141) = 5.40$, $p < .001$, $d = 0.45$]. Thus, participants correctly see prices as divvying up gains from trade between buyer and seller, but err in thinking that buyers often fail to benefit altogether (or even are made worse off).

Thus, the price manipulation did affect participants' judgments of welfare-change. But looking at Figure 4, the most striking finding is not the effect of price level, but the dramatic asymmetry between buyers and sellers at every level—regardless of price, buyers are much more likely to be perceived as non-beneficiaries compared to sellers. For example, at market price, buyers were deemed over six times likelier to be made worse-off by the transaction compared to sellers [$M = 0.61$, $SD = 0.76$ vs. $M = 0.09$, $SD = 0.33$ out of 2; $t(141) = 7.52$, $p < .001$, $d = 0.87$]. Indeed, even at 50% of the market price, buyers were thought nearly four times likelier to be made worse-off than sellers [$M = 0.50$, $SD = 0.73$ vs. $M = 0.13$, $SD = 0.40$; $t(141) = 5.21$, $p < .001$, $d = 0.62$]. Although the price-based trends suggest that people (correctly) believe that price makes a difference in allocating gains between buyer and seller, they frequently fail to perceive that buyers gain from trades, even at bargain prices.

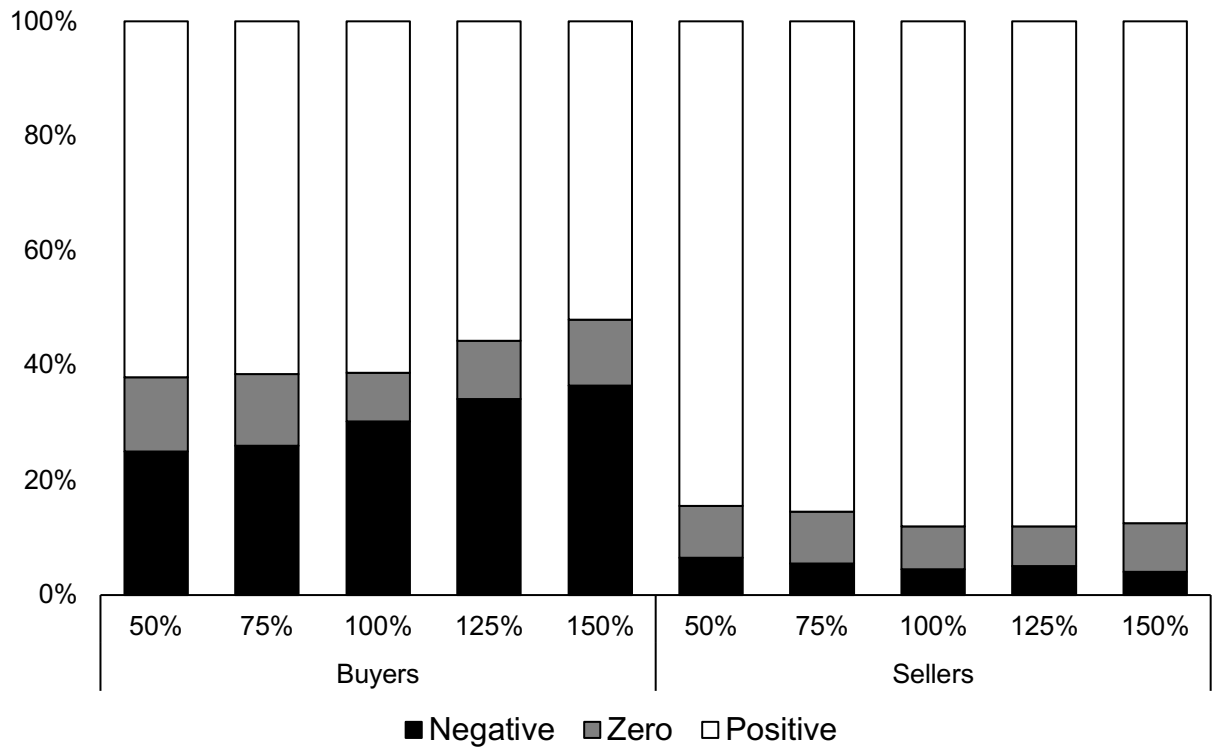

Figure 4. Proportion of transactions perceived as having negative, zero, or positive impact on buyers and sellers across prices (from 50% to 150% of typical retail price) in Study 4.

Discussion

Study 4 tested whether a "bargain" heuristic is a plausible explanation for zero-sum thinking—whether people tacitly substitute the question of relative welfare (before versus after a transaction) for relative bargain (compared to other transactions of similar goods and services). People do seem to calibrate their judgments of relative welfare somewhat depending on the price of the transaction—buyers are seen as somewhat less likely to lose from trade when the purchase price is 50% rather than 150% of the market price, and sellers somewhat less likely to benefit (though the latter trend was not significant). But the most striking finding is just how shallow these trends are: Nearly 40% of buyers were seen as failing to gain from trade even when the prices were 50% below market price—barely less win–win denial than in Study 1. Thus, while the "bargain" heuristic seems to play some role in win–win denial, it does not appear to be in the driver's seat.

General Discussion

Voluntary transactions benefit both parties—this is a truth universally acknowledged among economists. Here, we showed that non-economists have grave doubts about this truth. When evaluating the relative welfare of buyers and sellers (in monetary exchanges) and of traders (in barters), people frequently claimed that some parties to the transactions were worse off afterwards—in flagrant violation of the commonsense insight that both parties gain from voluntary trades. Buyers were much more likely to be thought worse-off than sellers and barters were frequently seen as failing to benefit

either party. Overall, the overwhelming majority of participants claimed that at least some of the parties did not benefit from one or more exchanges (see Study 1).

We also examined the mechanisms underlying this win–win denial, finding evidence in favor of two key mechanisms. First, across all studies buyers were consistently seen as less likely to benefit from exchange than sellers, and barters were often seen as not benefitting either party. This is consistent with *intuitive mercantilism*—the idea that a person's welfare is determined by their monetary wealth, not by their command of useful goods and services. Perceived benefit flows with currency, so that sellers are seen as better-off, buyers as worse-off, and traders as experiencing no change. Despite perennial attempts to conquer mercantilist thinking by economists (e.g., Bastiat, 2007/1845, Smith, 1999/1776), this sort of thinking may be so cognitively natural that even extensive economics education does not stamp it out. In our experiments, mercantilist thinking also manifested in a smaller degree of win–win denial when payments were described in terms of time rather than money (Study 2; see also Study S2 in the Supplementary Materials and Johnson & Park, 2019 for related results in the domain of charitable giving).

Second, win–win denial seems to be exacerbated by issues in our theory of mind. Specifically, people are naïve realists, making a perspective-taking error in which they interpret their own preferences as ground truth, neglecting that others have different preferences and reasons for their actions. Merely reminding people that the buyers and traders had reasons for their choices (even empty reasons such as "Mary wanted the chocolate bar") reduced the incidence of win–win denial (Study 3; see also Study S3 in the Supplementary Materials). Other results reported in the Supplementary Materials were also consistent with this idea. Making the preference of buyers and traders more salient reduced win–win denial (Study S4), as did asking participants to rate the parties' *perceived* gain or loss (Study S5). Together, these results suggest that people do not spontaneously reflect on the fact that parties to exchanges have reasons for their behavior, leading them to discount potential gains from trade.

We also addressed two other possible mechanisms. One possibility is *evolutionary mismatch*—that we were adapted well for exchange in our ancestral environment, but not the modern global economy (e.g., Pinker, 2003). One specific version of this proposal is that coalitional rivalry causes us to be averse to international trade (Boyer & Petersen, 2018)—hence the talk of "winners" and "losers." Weak versions of this approach are highly plausible—for instance, noting that people likely have poor intuitions about exchange in our complex global marketplace with complex and opaque chains of causation and with money-mediated transactions poorly suited to our adapted intuitions (Pinker, 2003). But some predictions of the evolutionary mismatch approach seem ill-equipped to explain our data. First, people should be readily perceive barters as positive-sum, since we are thought to be adapted for like-kind exchange (e.g., Cosmides & Tooby, 1992; Fiske, 1992). But we see, if anything, the opposite finding, with rampant zero-sum thinking about barters. Second, this account plausibly implies that our ancestors would have assigned value to goods but not to worthless pieces of paper, but this seems to predict the of our finding that buyers (who gain goods but give up money) are seen as worse-off than sellers. Finally, if coalitional rivalry accounts for aversion to trade, then we should see win–win denial at the level of countries but not of individuals, in contrast to the current results. It is, however, eminently possible that coalitional rivalry exacerbates at the international level the win–

win denial and zero-sum thinking we already see at the individual level (indeed, this is consistent with findings in Johnson, Zhang, & Keil, 2019). Thus, although we agree that our evolutionary past has not endowed us with the tools for understanding modern economic exchange, it is unclear that evolutionary mismatch is the main driver of zero-sum thinking (Johnson, 2018).

Another possible driver of win–win denial is a *"bargain" heuristic*, substituting the question of relative welfare (before and after the exchange) for bargain quality (relative to other deals). This would predict that buyers are seen as worse-off and buyers as better-off when the prices are high, but that buyers would be seen as better-off and sellers worse-off when the prices are low. In Study 4, we did find that people view sellers as benefitting at buyers' expense to a greater extent when prices are high, but even at bargain prices buyers were still often seen as worse-off, and much more frequently than sellers. Thus, although such heuristics might contribute to win–win denial, they do not seem to be the root cause.

Is Win–Win Denial Rational?

The conclusion that voluntary transactions benefit both parties rests on assumptions, and can therefore admit exceptions when these assumptions do not hold. Voluntary trades are mutually beneficial when the parties are performing rational, selfish cost–benefit calculations and when there are no critical asymmetries in information (e.g., fraud). There are several ways that violations of these assumptions could lead a transaction not to be win–win. Consumers could have inconsistent preferences over time, such that something believed to be beneficial at one time proves non-beneficial later on (e.g., liking a shirt when one buys it in the store, but growing weary of it after a couple months). Consumers could have self-control failures, making an impulse purchase that proved unwise in the longer term. Consumers could have other-regarding preferences, buying something that benefits someone else but not oneself. Finally, the consumer could be deceived by a seller who knows that the product will not satisfy their preferences (e.g., a crooked used-car salesman).

These are of course more than theoretical possibilities—many demonstrations of human irrationality have been demonstrated in lab and field studies (Frederick et al., 2009; Loewenstein & Prelec, 1992; Malmandier & Tate, 2005 among many others). The key question is whether the real-world prevalence of irrationality and fraud is sufficient to justify the conclusion that ordinary consumer transactions—like those tested here—are so riddled with incompetence that our participants were right to deny that transactions are typically win–win. We respond to this challenge with four points.

First, an empirical point. It is not just the magnitude of win–win denial of interest here, but how this magnitude responds to our experimental manipulations. It is hard to see how the effects of time-framing or cueing participants to buyers' reasons would produce the effects that they do, independent of the mechanisms we have proposed for win–win denial (namely mercantilism and theory of mind). It is especially difficult to see why people would claim that barters make *neither* party better-off if the issue is exploitation. Thus, even if the magnitude of the effects is reasonable in some conditions of some of our experiments because people's intuitions are attuned to the (allegedly) large extent of market failures, some of the patterns we see and the differences in these patterns across conditions seem to necessitate the mechanisms we propose.

Second, a sanity check. We tested intuitions about a range of typical consumer transactions in our items, finding consistent effects across items (see Part A of the Supplementary Materials). Is it really that plausible that people are impulsively hiring plumbers or that their hair stylists are routinely fraudsters? If such ordinary transactions are actually making consumers worse-off, it is very difficult to see how the rise of market economies has brought prosperity to much of the world—indeed, if win–win denial correctly describes most consumer transactions, one should predict a negative relationship between well-being and economic activity (contradicting the large association between subjective well-being and per capita income across countries; Stevenson & Wolfers, 2013). In our view, one can acknowledge occasional consumer irrationalities, while not thereby concluding that all or most market activity is irrational, which, we submit, would fly in the face both of economic science and common sense. Actually, to claim that consumers are consistently irrational threatens paradox: The more one thinks that consumers are irrational in general, the more one must believe that participants in the current experiments are (rationally) attuned to their own irrationality.

Third, a call for reflection. Economists tend to be much more accepting of the benefits of free markets compared to non-economists, and this is true even though economists tend to be to the political left of the general public (Caplan, 2002). If a (non-economist) reader's intuitions are aligned with our participants—believing that a very large proportion of ordinary consumer transactions are zero- or negative-sum—it is worth considering the possibility that this reader is herself falling prey to the same biases as other non-economists. (Indeed, Kahan et al., 2017, found that the most numerate individuals were the most skilled at distorting data to support their ideologically-derived conclusions.) This is not of course to say that there are no reasonable grounds for skepticism: simply that non-economist readers may want to consider the possibility that economists are genuine experts.

Fourth, a philosophical point. Our central project is an empirical one, not a normative one. We find that win–win denial is ubiquitous and we provided empirically-driven explanations of where it comes from. The proper interpretation of these results—as grievous error or knowing wisdom—must be situated relative to the reader's own view of the ground truth on these issues. If, reaching the end of this article, the reader believes that markets are indeed in perpetual failure due to human frailty, it is still useful to know that people are (perhaps paradoxically) not so frail that they fail to notice this. The purpose of this article would be, then, documenting a *belief* rather than a *mistake*.

Explanations of Intuitive Mercantilism

Although we have identified intuitive mercantilism as one of the underlying mechanisms of zero-sum thinking, we have not explained where mercantilism itself comes from: Why is this belief system, as opposed to Smithian political economy, psychologically intuitive? We do not provide direct evidence in this article, but here we consider three possible (and not mutually exclusive) theories—a *heuristic for value*, a *competition for relative rank*, and a *self-control mechanism*. These correspond roughly to the classic functions of money (Jevons, 1875)—as a *unit of account*, *medium of exchange*, and *store of value*.

Heuristic for value. According to mainstream microeconomics, consumers purchase a product when its subjective value (the utility the consumer can expect to gain) is greater than its opportunity cost (the value of the next-most-desirable alternative). These are situations that produce consumer

surplus (e.g., Mankiw, 2017) and are therefore win–win. We need not assume that consumers are explicitly performing such calculations when purchasing products. But in our task—and any real-world situation in which people deliberate whether a transaction makes two parties better-off or worse-off—we *are* asking people to calculate consumer surplus, the difference between value and opportunity cost.

We speculate that people use objective and salient *monetary* costs, rather than subjective values, as heuristics for valuation. Value estimations are difficult—for example, although people can make consistent *relative* value judgments, their absolute value judgments are all over the map (e.g., Ariely, Loewenstein, & Prelec, 2003). Substituting monetary costs for subjective value is likely attractive due to money's function as a *unit of account*, used to denominate transactions and facilitate comparability of value across goods. Because each dollar is equivalent to every other dollar, it is useful to express the value of goods in monetary terms even when bartering. Ultimately, value in economic theory is imbued in a *subjective* currency of personal utility, but such private and subjective notions are necessarily more nebulous than the socially shared system of prices, denominated in an objective, monetary currency. It is thus plausible that monetary costs would take the role of subjective value in much economic discourse. Yet, this is a fundamental mistake. Since consumer surplus simply *is* the difference between a product's subjective value and its (monetary) cost, this heuristic leads to the conclusion that positive consumer surplus is impossible!

How does this account of mercantilism relate to the other mechanisms we documented in this article? It is linked to theory-of-mind errors as a manifestation of the "physical fallacy" described by economists (Sowell, 1980)—the idea that goods have precisely one value at a given time, when value is in fact subjective and differs across people and across time. In our experiments, we found that people often fail to think about buyers' subjective reasons and preferences as drivers of (utility-increasing behaviors), instead appearing to substitute their own judgments of value. The idea that people use monetary costs as a heuristic for value suggests that these judgments of value are themselves tied unduly to these monetary costs. Thus, a mercantilist heuristic for value is one particular way to fill in the gap left by this perspective-taking error, but not the only logical possibility.

It is also linked with the bargain heuristic, but these are subtly different. The bargain heuristic substitutes judgments of welfare (comparing utility after an exchange to *before the exchange*) for judgments of bargain quality (comparing utility after an exchange to *other counterfactual exchange opportunities*). This need not be tied to money at all—in Study 4 we manipulated the exchange's quality compared to other exchanges by altering the price, but this heuristic could apply as easily to barters. In contrast, the mercantilist heuristic for value substitutes judgments of utility or subjective value for monetary costs or prices, as a common (albeit problematic) yardstick for value.

Competition for relative rank. Economists from Smith onwards conceptualize markets as positive-sum because gains from trade and economic growth can lead to more consumption for everyone. But we may be more plausibly evolved for Darwinian rather than Smithian competition. In Darwinian competitions, it is our relative standing that matters, as in competitions for dominance within a group or for mating. Indeed, there is evidence that people are often more interested in their *relative* standing rather than their *absolute* level of consumption (e.g., Boyce, Brown, & Moore, 2010; Clark, Frijters, & Shields, 2008). Frank (2005, 2012) argues that these two forms of consumption co-

exist within the modern economy, with the utility of some goods depend on the absolute consumption level (e.g., vacation time) and others on relative consumption (e.g., house size).

This idea by itself does not explain mercantilism, since both absolute and relative consumption levels are raised by increasing consumption, although some goods (e.g., expensive clothing or cars) may be more useful in the relative status competition than others (Veblen, 1899). However, in some status competitions it may not be relative actual or *realized* consumption, but relative *potential* consumption that matters. For example, Cold War era anxiety over economic growth in the Soviet Union was not motivated by jealousy over the perceived luxuries with which Soviet citizens lived (*realized consumption*), but out of fear of the military power that economic growth could bring (*potential consumption*). Likewise, it serves the interest of a potential romantic partner to seek out mates with more resources (*potential consumption*) rather than spendthrift consumption in the present. Indeed, to the extent that conspicuous consumption sends a positive signal, it is because it demonstrates that resources are *so abundant* that even wholly useless consumptive acts are painless (Zahavi, 1975).

Money is a measure of potential consumption because it is a *medium of exchange*—it can be exchanged for many different consumption goods. If you need to buy a house, pay your child's university tuition, or start a war with another country, it is easier to do so if you store your wealth in dollars than in cans of lima beans, because dollars can be more readily converted into houses, tuition, munitions—or indeed lima beans—than the reverse. Hence, to the extent that people are interested in relative standing rather than absolute consumption, and to the extent that relative standing is measured in *potential consumption* rather than *realized consumption*, people would privilege stored, money-denominated wealth over consumption in their notion of value. Thus, accumulating money can take on value not as a way to facilitate consumption, but as an end in itself. Intuitive as this may seem, it actually turns standard economics on its head: Savings and investment in standard economics are motivated by the desire to defer or smooth consumption over one's lifetime, not as ends in themselves (Friedman, 1957; Modigliani & Brumberg, 1954).

Self-control mechanism. Theories of mental accounting have documented many normative errors in how people think about money, for instance ignoring the fungibility of money across different kinds of mental accounts (e.g., Thaler, 1985). For example, people often hold money in low-interest savings account at the same time that they have higher-interest debts. One explanation for such behaviors is that they are self-control mechanisms—putting money into a savings account each month ensures that it is not spent, and hence may actually lead to better outcomes in the long-run (assuming that the consumer correctly judges that spending this money is not in their overall interest).

Mercantilism is a particularly deep kind of error, not in our mental accounting of money but in our mental accounting of subjective value or utility, in *substituting* money for these more nebulous, subjective notions. In systematically underestimating consumer surplus, mercantilism may help us to restrict our own spending, preserving resources that are liquid and storable over time. It would therefore be interesting to test whether mercantilist intuitions to be strongest among individuals high in self-control or propensity to plan (Lynch, Netemeyer, Spiller, & Zammit, 2010).

Future Directions

Although this article has identified the phenomenon of win–win denial and documented the contributions of several mechanisms, there is still much to learn about this phenomenon, its causes and consequences, and its boundary conditions.

First, a further contributing factor may be beliefs about sellers' power to dupe buyers (Bhattacharjee et al., 2017; Vohs, Baumeister, & Chin, 2007), for instance due to sellers' greater knowledge about the products. It is unclear to what extent this can account for the results of the current studies, since if anything participants were more prone to believing that buyers were made worse-off by purchases of goods (e.g., olive oil, where sellers typically have little expertise) than services (e.g., plumbing, where sellers typically have high expertise; see Part A of the Supplementary Materials for item-level results). Still, beliefs about information asymmetries could account for some portion of win–win denial and would make several predictions, for instance about how participants' judgments would shift when buyer vs. seller expertise or motives are manipulated, for product categories that lend versus do not lend themselves to expertise, or when information about the product or seller's reputation are available. For example, in situations where the buyer has more product knowledge than the seller, the typical mercantilist pattern should be attenuated or reversed to the extent that perceived information asymmetries drive win–win denial.

Second, a related reason participants might deny that buyers gain from trade would be denial that the buyers actually know their own interests—a belief associated with hard paternalism or the desire to block transactions for the consumer's own good even when the transaction is voluntary and the consumer is knowledgeable (Mill, 1859). The finding that participants are less prone to thinking that the consumer *believes* herself to be worse-off rather than *actually* worse-off (Study S5A) is consistent with this possibility. If this is the case, then the alignment between the buyers' and participants' conception of the buyer's interest would be a key moderating variable on win–win denial. In extreme cases, participants who believe that many people do not know their own interests might even claim that "transactions" without a counterparty (e.g., Robinson Crusoe deciding between fish and game for dinner) are welfare-depleting. Alternatively, people might rely on different mechanisms entirely for contemplating such non-social "transactions with nature."

A specific version of this idea is that asymmetries between buyers and sellers could result from differential beliefs about anticipated regret (Loomes & Sugden, 1982). Perhaps perhaps people believe that buyers have poor self-control whereas sellers do not have this problem, leading to frequent buyer's remorse but rare seller's remorse. Alternatively, if a transaction is seen as a risky gamble, then even if the gamble has positive expected utility (because the possible gains are likelier or more valuable than the possible losses), loss aversion may lead the possibility of anticipated regret to outweigh the possibility of anticipated satisfaction. This could lead to negative expected "emotional utility" even if the transaction is likely to be materially beneficial; yet sellers should not experience regret, since loss aversion is not found for goods given up "as intended" (Novemsky & Kahneman, 2005). However, this explanation is somewhat at odds with our experimental results: If sophisticated inferences about future mental states underlie beliefs that buyers lose from their transactions, then cueing participants to think about buyers' mental states should make these beliefs even stronger, contrary to our findings. Moreover, although sellers should not experience loss aversion for goods given up "as intended," this

logic also applies to buyers, since money is given up as intended in these exchanges—there is no loss aversion for money (Novemsky & Kahneman, 2005).

Third, participants beliefs about firms' costs may contribute to the belief that sellers are especially likely to benefit from transactions. Prior work on price fairness has found that consumers have dubious beliefs about firms' cost structures, for instance neglecting the role of inflation and attributing price differences across retailers to profits rather than costs (Bolton, Warlop, & Alba, 2003). If consumers neglect many of the costs sellers face and their associated benefits (e.g., convenience, speed, and customer service), they may overestimate sellers' profits, in turn triggering resentment and zero-sum thinking (Bhattacharjee et al., 2017). This may fit with a broader neglect of the benefits provided by economic middlemen who sit between producers and consumers (Bastiat, 2007/1850). One direction for future work would be to test how sellers' cost structure impacts perceptions both of firm profitability and zero-sum thinking. For example, when fixed costs are high, competition can drive the price below the firm's average cost—would people believe that firms are benefitting at buyers' expense in this situation? Or even if the firm was demonstrably selling below marginal cost?

Fourth, people may be more or less prone to win–win denial depending on the competitive environment in which the seller sits. According to standard economic theory, firms can extract a larger portion of the surplus value created by an exchange when they have few competitors. For example, a recent report from the US Food and Drug Administration analyzing the effects of deregulation in the generic drug market found that generic drugs manufactured by only one firm (i.e., a monopoly) are priced on average around 69% the price of the name-brand drug, with prices declining as the number of manufacturers increases, to 56% (for 2 producers), 44%, 27%, 18%, and 10% (for 6 producers), to less than 5% of the name-brand drug price for more than 10 producers (Conrad & Lutter, 2019).

Thus, manipulating the market structure (number of competitors) and other sources of market power (e.g., threat of new entrants; Porter, 1979) , and availability of market prices from competitors may all shape the extent of win–win denial. We would expect win–win denial to be stronger in more monopolistic markets and in competitive markets when the seller's price is high relative to competitors (see Bolton et al., 2003). Economic theory says that even in such cases, win–win denial is a mistake because consumer still benefit (albeit to a smaller degree) from individual transactions as they would not make the exchange if it produced negative marginal utility. However, we also suspect that most people overestimate the prevalence of true monopolies and neglect the threat of new entrants or substitute products, further exacerbating this problem. It would be particularly striking if people continue to believe that firms benefit at the expense of consumers under highly competitive conditions, as economic theory says that profits are competed away under perfect competition, leading to no net gain on the part of sellers (Mas-Colell et al., 1995).

Fifth, the framing of the exchanges may exacerbate win–win denial. Mental accounting theory (Thaler, 1985) distinguishes between the utility of consumption (i.e., eating the chocolate bar or driving the car one has bought) versus the utility of the transaction itself (i.e., giving up money in exchange for the product). Perhaps asking participants about who benefits from exchanges leads to a focus on transaction utility at the expense of consumption utility. This is consistent with the finding in Study 4 that better "bargains" (i.e., high in transaction utility) are perceived as at least somewhat more beneficial to buyers. Moreover, whereas the buyer's consumption utility will occur in the future, the

seller's economic gains from the trade itself occur immediately. In that case, focusing attention on the consumption experience itself may lead people to more readily recognize that transactions tend to benefit buyers as well as sellers.

Sixth, these results document a general bias in economic thinking, but it is one with applications to many economic policy issues. Anti-immigrant sentiment may be driven in part by the belief that there is a fixed number of jobs that are "taken" by immigrants (Esses et al., 2001; Louis et al., 2013), neglecting that immigrants are not only workers but also consumers. Similarly, the anti-trade sentiment recently seen in U.S. politics seems to be driven by the belief that other countries are "winning" at trade, whereas the U.S. is "losing." In a separate project, we demonstrate that win–win denial about international trade is, if anything, even more rampant than the effects we document in this article (Johnson et al., 2019). Understanding more broadly how win–win denial affects voters' attitudes may be valuable.

Seventh, further work into the psychological origins of mercantilism would be valuable both in basic scientific value and in potential applications. For example, understanding the reasons why currency is over-valued may be useful for understanding the psychological differences among various currencies and payment methods (e.g., cash versus credit cards) and for the design of new currencies, such as cryptocurrencies.

Finally, whereas our existing data and speculations about potential moderators have been driven by theoretical concerns about underlying mechanisms, an important practical project is understanding how to frame economics in a way that minimizes errors in economic thinking. Thus, seeing how these manipulations work in conjunction, and how they might be exported to more real-world rather than laboratory conditions, can be useful for both policy and economics education.

Further Puzzles

These results partly resolve one puzzle—when and why people believe that economic transactions are zero-sum—but raise several further puzzles when taken together with the broader literature in behavioral and traditional economics.

First, if people really believed that so many consumer transactions are zero-sum, then why doesn't economic activity grind to a halt? One possibility could be that people have more accurate views that their own economic transactions are positive-sum, with win–win denial only a powerful force for evaluating hypothetical or third-person transactions such as those we study here. After all, we find that better perspective-taking reduces win–win denial, and surely we are capable of taking our *own* perspectives? Perhaps surprisingly, we find in a separate project (Johnson, Zhang, & Keil, 2021) that consumers often claim that their own past transactions make them either worse-off or no better off, and even make similar claims about planned *future* transactions. Thus, there appears to be a striking attitude–behavior gap here: Whereas people's lay theories of exchange seem to produce strong intuitions that consumers are often made worse-off by their purchases, these attitudes do not seem to manifest (in most cases, fortunately) in their actions. Perhaps this gap is driven by differences in what is considered relevant when evaluating exchanges more abstractly from a distance versus more concretely from a nearby temporal perspective (Trope & Liberman, 2010), with the latter conditions

prompting more thoughts about the consumption experience itself (see "Future Directions" above). In any case, we think this is a genuine puzzle deserving of further research.

Second—and exacerbating the first puzzle—some transactions really *are* zero- or negative-sum, yet people make them anyway (Camerer, 2003). For example, poker games are zero-sum in that the winner's gains exactly balance out the other players' losses. Given that people have some private information about their skill, players unlikely to win should be unwilling to play since the expected value of playing is negative. Of greater economic consequence, ordinary investors making short-term trades in the hope of beating the market are engaging in zero-sum trades while neglecting the (likely greater) information that the party on the other end of the transaction has. Such willingness to take zero-sum bets is known as the *Groucho Marx theorem*, after the comedian's observation that he would never join a club willing to take him as a member. Thus, the puzzle that consumers are willing to make positive-sum transactions despite judging them to be zero-sum becomes even more mysterious, as they will even make *genuinely* zero- or negative-sum transactions. One important question for future research is whether people can recognize the distinction between positive- and zero-sum transactions. If people treat all economic transactions as zero-sum by default, yet learn that frequent transactions are required to gain life's necessities, then perhaps people become numbed to these genuinely zero-sum transactions and do not adequately appreciate their drawbacks. Further, zero-sum transactions like day-trading are made under uncertainty and may be especially susceptible to an illusion of skill.

Third, these results are in seeming tenson with demonstrations that people tend to overestimate how much others are willing to pay (Frederick, 2012) and how strong other's preferences are (Jung, Moon, & Nelson, 2020). The results of Study 3 seem to suggest that people do not spontaneously recognize that willingness-to-pay is *evidence* of a strong preference, without cueing to consider that people have reasons for their choices. One possible interpretation of this body of evidence is that, despite believing that others' preferences are stronger than their own, they supply their *own* preferences when evaluating the rationality of others' exchanges and only recognize the relevance of others' preferences when cued to do so. However, this interpretation is speculative and requires further study.

Fourth, whereas economists insist that the true cost of something is the next-best-alternative that is given up (its *opportunity cost*), people often seem to neglect opportunity costs. For example, when a choice is framed as buying a product versus not buying it, people are more likely to make the purchase than when the choice is instead framed as buying a product versus keeping its cost for other purchases (Frederick et al., 2009). Subsequent research (Spiller, 2011) has therefore examined when opportunity costs are neglected versus considered (e.g., when resource constraints are more salient) and who is most likely to consider them (e.g., consumers high in propensity to plan). In Study 4, people were more likely to deny that buyers benefited from transactions when prices were high, relative to implicit market prices. This demonstrates that people naturally account for reference prices even when they are not provided explicitly and even when they are not normatively relevant to the task. This form of opportunity cost consideration may merit further investigation.

Folk Economics

Since the beginning of professional economics, economists have complained bitterly about people's economic ignorance. Recent survey data comparing the views of economists and laypeople

suggests that little progress has been made since the time of Adam Smith—even though economic science has advanced greatly, our intuitive theories seem to be stuck in time (Caplan, 2007). The ubiquity of zero-sum language among politicians from the Elizabethan era until today betray a depressing lack of progress in economic understanding.

We view the current studies as a step toward a systematic study of people's intuitive theories of economics (Boyer & Petersen, 2018; Leiser & Shemesh, 2018). Humans have sophisticated intuitive theories of intuitive psychology (Apperly, 2010), physics (Carey, 2009), and biology (Keil, 1994), and although these intuitive theories often diverge from scientific theories (Shtulman, 2017), humans seem to have some innate capacities to understand these domains in an adaptive manner. But to the extent that innate ideas influence our intuitive economics, these innate ideas come from a radically different evolutionary past—the principles of physics, biology, and psychology have not changed much over the millennia, but the principles of economics certainly have. Modern economies are mediated by money rather than barter; they are far more specialized and globalized; and as a result we typically trade with people we do not know. If we have any innate economic intuitions, they developed in a world of barter with close acquaintances in a much less specialized economy. Our modern and ancestral economies both depended largely on trust, but many of the cues that trigger trust are missing from modern markets. One possibility is that, lacking an autonomous set of folk-economic intuitions, we instead rely on domain-specific folk-psychological intuitions, such as a bias for intentional explanations (Johnson & Nagatsu, 2021; Leiser et al., 2010; Rosset & Kelemen, 2008).

Mapping our intuitive economic concepts and theories—whether they are best accounted for in evolutionary or heuristic terms—is an important project for ongoing research. Despite the nascent state of folk economics, there has been a recent flurry of research into this topic (e.g., Leiser & Shemesh, 2018). Boyer and Petersen (2018) review a number of folk-economic beliefs and analyze these beliefs in terms of an evolutionary framework. Although we are not bullish on the prospects of evolutionary psychology to explain zero-sum thinking or aversion to trade for the reasons given above, we are enthusiastic about this line of inquiry more generally. For example, laypeople are known to be far more skeptical about the economic impact of immigration compared to economists (Caplan, 2007), and evolutionary accounts in terms of coalitional rivalry are well-suited to explaining this difference.

Practical Implications

Democracy involves a trade-off—political leaders must be responsive to people's expressed interests, limiting the range of potential self-interested choices they can make (Bueno de Mesquita, 2003). But the policies we get in place of dictatorship will not be effective if people do not in fact know what is in their interest (Brennan, 2016; Caplan, 2007). Thus, if democracy is to be an effective institution for maximizing everyone's well-being, it is critical that voters be informed not only about the narrow issues of the day, but perhaps more importantly about the fundamental principles governing the economy.

No one knows what populist politicians really think about trade. Perhaps they do not really believe that it is zero-sum. But they surely say so, and the current research shows, regrettably, that they have good game-theoretic reasons for it: Win–win denial is a pervasive element in human psychology, and is equally prevalent on the political left and right (see Part C in the Supplementary Materials). Their

ability to harness this populist sentiment is surely one source of their political power, with all the policy consequences that entails. Moreover, in a separate set of studies, we find that mercantilism plays a powerful role in anti-trade sentiments as well, and can lead international trade to be deemed immoral (Johnson et al., 2019).

We may not know how to solve public policy, but at least we now have a good idea about why many people oppose free trade and open immigration: Not only do trade and immigration harness in-group bias (see Caplan, 2007), but they also violate the logic of a zero-sum game—if we allow China and Mexico to get part of the pie, then the part left for us will be smaller. The moderating role of economics knowledge may be one source of encouragement, and highlights the importance of emphasizing basic principles in economics courses, without losing students in mathematical formalisms. But perhaps the experimental moderators uncovered here—interventions that emphasize the mental states of buyers, but de-emphasize currency—are the most promising route to attenuating zero-sum thinking in the classroom and in the political arena.

References

Acemoglu, D., & Robinson, J. A. (2012). *Why nations fail: The origins of power, prosperity, and poverty*. New York, NY: Crown.

Apperly, I. A. (2010). *Mindreaders: The cognitive basis of "theory of mind."* Abingdon, UK: Routledge.

Ariely, D., Loewenstein, G., & Prelec, D. (2003). "Coherent arbitrariness": Stable demand curves without stable preferences. *Quarterly Journal of Economics*, *118*, 73–106.

Baron, J., & Kemp, S. (2004). Support for trade restrictions, attitudes, and understanding of comparative advantage. *Journal of Economic Psychology*, *25*, 565–580.

Bastiat, F. (2007). *Economic sophisms: First series*. In *The Bastiat Collection* (Trans. P. J. Stirling.) Auburn, AL: Ludwig von Mises Institute. (Original work published 1845.)

Bastiat, F. (2007). What is seen and what is not seen. In *The Bastiat Collection* (Trans. P. J. Stirling.) Auburn, AL: Ludwig von Mises Institute. (Original work published 1850.)

Bazerman, M. H., Baron, J., & Shonk, K. (2001). *You can't enlarge the pie: The psychology of ineffective government*. New York: Basic Books.

Bazerman, M. H., & Neale, M. (1983). Heuristics in negotiation: Limitations to dispute resolution effectiveness. In M. H. Bazerman, & R. Lewicki (Eds.), *Negotiations in organizations* (pp. 51–67). Beverly Hills, CA: Sage.

Becker, G. S. (1962). Irrational behavior and economic theory. *Journal of Political Economy*, *70*, 1–13.

Bhattacharjee, A., Dana, J., & Baron, J. (2017). Anti-profit beliefs: How people neglect the societal benefits of profit. *Journal of Personality and Social Psychology*, *113*, 671–696.

Blackwell, R. D., Miniard, P. W., & Engel, J. F. (2001). *Consumer behavior* (9th Ed). Ft. Worth, TX: Harcourt.

Blythe, J. (2013). *Consumer behaviour* (2nd Ed.). Thousand Oaks, CA: Sage.

Bolton, L. E., Warlop, L., & Alba, J .W. (2003). Consumer percpetions of price (un)fairness. *Journal of Consumer Research*, *29*, 474–491.

Boyce, C. J., Brown, G. D. A., & Moore, S. C. (2010). Money and happiness: Rank of income, not income, affects life satisfaction. *Psychological Science*, *21*, 471–475.

Boyer, P., & Petersen, M. B. (2018). Folk-economic beliefs: An evolutionary cognitive model. *Behavioral and Brain Sciences*, *41*, e158.

Brennan, J. (2016). *Against democracy*. Princeton, NJ: Princeton University Press.

Bueno de Mesquita, B. (2003). *The logic of political survival*. Cambridge, MA: MIT Press.

Burleigh, T. J., Rubel, A. N., & Meegan, D. V. (2017). Wanting 'the whole loaf': Zero-sum thinking about love is associated with prejudice against consensual non-monogamists. *Psychology & Sexuality*, *8*, 24–40.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton, NJ: Princeton University Press.

Caplan, B. (2002). Systematically biased beliefs about economics: Robust evidence of judgemental anomalies from the survey of Americans and economists on the economy. *Economic Journal*, *112*, 433–458.

Caplan, B. (2007). *The myth of the rational voter: Why democracies choose bad policies*. Princeton, NJ: Princeton University Press.

Carey, S. (2009). *The origin of concepts*. Cambridge, MA: Harvard University Press.

Clark, A. E., Frijters, P., & Shields, M. A. (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. *Journal of Economic Literature*, *46*, 95–144.

Conrad, R., & Lutter, R. (2019). Generic competition and drug prices: New evidence linking greater generic competition and lower generic drug prices. *FDA Report*.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31*, 187–276.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. H. Barkow (Ed.), *The Adapted Mind* (pp. 163–228). New York, NY: Oxford University Press.

Davidai, S., & Ongis, M. (2019). The politics of zero-sum thinking: The relationship between political ideology and the belief that life is a zero-sum game. *Science Advances*, *5*, eaay3761.

Davis, M. H., Conklin, L., Smith, A., & Luce, C. (1996). Effect of perspective taking on the cognitive representation of persons: A merging of self and other. *Journal of Personality and Social Psychology, 70*, 713–726.

De Dreu, C. K. W., Koole, S. L., & Steinel, W. (2000). Unfixing the fixed pie: A motivated information-processing approach to integrative negotiation. *Journal of Personality and Social Psychology*, *79*, 975–987.

De Freitas, J., & Johnson, S. G. B. (2018). Optimality bias in moral judgment. *Journal of Experimental Social Psychology*, *79*, 149–163.

Drucker, P. F. (1954). *The practice of management*. New York, NY: Harper & Row.

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology*, *57*, 1082–1090.

Epley, N., Keysar, B., & Van Boven, L. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology*, *87*, 327–339.

Esses, V. M., Dovidio, J. F., Jackson, L. M., & Armstrong, T. L. (2001). The immigration dilemma: The role of perceived group competition, ethnic prejudice, and national identity. *Journal of Social Issues*, *57*, 389–412.

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.

Fisher, R., Ury, W., & Patton, B. (2011). *Getting to yes: Negotiating agreement without giving in*. London, UK: Penguin. (Original work published 1981.)

Fiske, A. P. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review*, *99*, 689–723.

Frank, R. H. (2005). Positional externalities cause large and preventable welfare losses. *American Economic Review*, *95*, 137–141.

Frank, R. H. (2012). *The Darwin economy: Liberty, competition, and the common good*. Princeton, NJ: Princeton University Press.

Frederick, S. (2012). Overestimating others' willingness to pay. *Journal of Consumer Research*, *39*, 1–21.

Frederick, S., Novemsky, N., Wang, J., Dhar, R., & Nowlis, S. (2009). Opportunity cost neglect. *Journal of Consumer Research*, *36*, 553–561.

Friedman, M. (1957). *A theory of the consumption function*. Princeton, NJ: Princeton University Press.

Friedman, M. (1962). *Capitalism and freedom*. Chicago, IL: University of Chicago Press.

Hubbard, T. L. (1996). Representational momentum, centripetal force, and curvilinear impetus. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1049–1060.

Hume, D. (1985). On the balance of trade. In *Essays Moral, Political, and Literary* (pp. 308–326). Indianapolis, IN: Liberty Fund. (Original work published 1752.)

IGM Forum (2012). Free trade. *IGM Economic Experts Panel.*

Jevons, W. S. (1875). *Money and the mechanism of exchange*. Boston, MA: D. Appleton and Co.

Johnson, S. G. B. (2018). Why do people believe in a zero-sum economy? *Behavioral and Brain Sciences*, *41*, e172.

Johnson, S. G. B., & Nagatsu, M. (2021). Structural and psychological explanation in folk and scientific economics. *Working paper.*

Johnson, S. G. B., & Park, S. Y. (2019). *Moral signaling through donations of money and time*. Available at PsyArXiv: https://psyarxiv.com/tg9xs/.

Johnson, S. G. B., & Rips, L. J. (2014). Predicting behavior from the world: Naïve behaviorism in lay decision theory. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 695–700). Austin, TX: Cognitive Science Society.

Johnson, S. G. B., & Rips. L. J. (2015). Do the right thing: The assumption of optimality in lay decision theory and causal judgment. *Cognitive Psychology*, *77*, 42–76.

Johnson, S. G. B., Zhang, J., & Keil, F. C. (2018). Psychological underpinnings of zero-sum thinking. In T. T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society* (pp. 566–571). Austin, TX: Cognitive Science Society.

Johnson, S. G. B., Zhang, J., & Keil, F.C. (2019). *Consumers' beliefs about the effects of trade*. Available at SSRN: https://papers.ssrn.com/abstract=3376248.

Johnson, S. G. B., Zhang, J., & Keil, F.C. (2021). *Zero-sum thinking in self-perceptions of consumer welfare.* Working paper.

Jung, M. H., Moon, A., & Nelson, L. D. (2020). Overestimating the valuations and preferences of others. *Journal of Experimental Psychology: General, 149*, 1193–1214.

Kagel, J. H., Battalio, R. C., Rachlin, H., Green, L., Basmann, R. L., & Klemm, W. R. (1975). Experimental studies of consumer demand behavior using laboratory animals. *Economic Inquiry*, *13*, 22–38.

Kahan, D. M., Peters, E., Dawson, E. C., & Slovic, P. (2017). Motivated numeracy and enlightened self-government. *Behavioural Public Policy*, *1*, 54–86.

Kahneman, D. (2002). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, *93*, 1449–1475.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive thought* (pp. 49–81). New York, NY: Cambridge University Press.

Keil, F. C. (1994). The birth and nurturance of concepts by domains: The origins of concepts of living things. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 234–254). New York, NY: Cambridge University Press.

Khon, Z., Johnson, S. G. B., & Hang, H. (2020). *Lay theories of manipulation: Do consumers believe they are susceptible to marketers' trickery?* Available at PsyArXiv: https://psyarxiv.com/8x63c/.

Koehler, J. J. (1996). The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral & Brain Sciences*, *19*, 1–53.

Krugman, P. (1996). *Pop internationalism*. Cambridge, MA: MIT Press.

Langer, E., Blank, A., & Chanowitz, B. (1978). The mindlessness of ostensibly thoughtful action: The role of "placebic" information in interpersonal interaction. *Journal of Personality and Social Psychology*, *36*, 635–642.

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, *92*, 805–824.

Leiser, D., Bourgeois-Gironde, S., & Benita, R. (2010). Human foibles or systemic failure: Lay perceptions of the 2008–2009 financial crisis. *Journal of Socio-Economics*, *39*, 132–141.

Leiser, D., & Shemesh, Y. (2018). *How we misunderstand economics and why it matters: The psychology of bias, distortion, and conspiracy*. Abingdon, UK: Routledge.

Loewenstein, G., & Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *Quarterly Journal of Economics*, *107*, 573–597.

Louis, W. R., Esses, V. M., & Lalonde, R. N. (2013). National identification, perceived threat, and dehumanization as antecedents of negative attitudes toward immigrants in Australia and Canada. *Journal of Applied Social Psychology*, *43*, E156-E165.

Lynch, J. G., Netemeyer, R. G., Spiller, S. A., & Zammit, A. (2010). A generalizable scale of propensity to plan: The long and the short of planning for time and money. *Journal of Consumer Research*, *37*, 108–28.

Mai, R., Hoffmann, S., Lasarov, W., & Buhs, A. (2017). Ethical products = less strong: How explicit and implicit reliance on the lay theory affects consumption behaviors. *Journal of Business Ethics*. Advanced online publication.

Malmandier, U., & Tate, G. (2005). CEO overconfidence and corporate investment. *Journal of Finance*, *60*, 2661–2700.

Mankiw, N. G. (2017). *Economics* (6th Ed.). Boston, MA: Cengage.

Marshall, A. (1949). *The pure theory of foreign trade: The pure theory of domestic values*. London, UK: London School of Economics and Political Science. (Original work published 1879.)

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic Theory*. Oxford, UK: Oxford University Press.

Meegan, D. V. (2010). Zero-sum bias: Perceived competition despite unlimited resources. *Frontiers in Psychology*, *1*, 191.

Mill, J. S. (1859). *On liberty*. London, UK: Longman, Roberts, & Green.

Modigliani, F., & Brumberg, R. H. (1954). Utility analysis and the consumption function: an interpretation of cross-section data. In K. K. Kurihara (Ed.), *Post-Keynesian economics* (pp. 388–436). New Brunswick, NJ: Rutgers University Press.

Newman, G. E., Gorlin, M., & Dhar, R. (2014). When going green backfires: How firm intentions shape the evaluation of socially beneficial product enhancements. *Journal of Consumer Research*, *41*, 823–839.

Novemsky, N. ,& Kahneman, D. (2005). The boundaries of loss aversion. *Journal of Marketing Research*, *42*, 119–128.

Pilditch, T. D., Fenton, N., & Lagnado, D. (2019). The zero-sum fallacy in evidence evaluation. *Psychological Science*, *30*, 250–260.

Pinker, S. (2003). *The blank slate: The modern denial of human nature*. New York, NY: Viking.

Polanyi, K. (1944). *The great transformation: The political and economic origins of our time*. New York, NY: Rinehart.

Porter, M. E. (1979). How competitive forces shape strategy. *Harvard Business Review*, *57*, 137–145.

Pronin, E., & Kugler. M. B. (2010). People believe they have more free will than others. *Proceedings of the National Academy of Sciences, 107*, 22469–22474.

Ricardo, D. (2004). *The principles of political economy and taxation*. Mineola, NY: Dover. (Original work published 1817.)

Ries, A., & Trout, J. (1981). *Positioning: The battle for your mind*. New York, NY: McGraw-Hill.

Ross, L., Greene, D., & House, P. (1977). The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, *13*, 279–301.

Ross, L., & Ward, A. (1997). Naive realism in everyday life: Implications for social conflict and misunderstanding. In E. S. Reed, E. Turiel, & T. Brown (Eds.), *Values and knowledge* (pp. 103–135). Mahwah, NJ: Erlbaum.

Rosset, E. (2008). It's no accident: Our bias for intentional explanations. *Cognition*, *108*, 771–780.

Różycka-Tran, J., Boski, P., & Wojciszke, B. (2015). Belief in a zero-sum game as a social axiom: A 37-nation study. *Journal of Cross-Cultural Psychology*, *46*, 525–548.

Rubin, P. (2003). Folk economics. *Southern Economic Journal*, *70*, 157–171.

Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological Bulletin, 134*, 207–222.

Shtulman, A. (2017). *Scienceblind: Why our intuitive theories about the world are so often wrong*. New York, NY: Basic Books.

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, *69*, 99–118.

Smith, A. (1999). *An inquiry into the nature and causes of the wealth of nations*. London, UK: Penguin. (Original work published 1776.)

Sowell, T. (1980). *Knowledge and decisions*. New York, NY: Basic Books.

Sowell, T. (2008). *Economic facts and fallacies*. New York, NY: Basic Books.

Stevenson, B., & Wolfers, J. (2013). Subjective well-bieng and income: Is there any evidence of satiation? *American Economic Review: Papers and Prcoeedings*, *103*, 598–604.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, *1*, 39–60.

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science*, *4*, 199–214.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology, 46*, 35–57.

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review, 117*, 440–463.

Veblen, T. (1899). *A theory of the leisure class*. New York, NY: Macmillan.

Vohs, K. D., Baumeister, R. F., & Chin, J. (2007). Feeling duped: Emotional, motivational, and cognitive aspects of being exploited by others. *Review of General Psychology*, *11*, 127–141.

Zahavi, A. (1975). Mate selection—A selection for a handicap. *Journal of Theoretical Biology*, *53*, 205–214.

Win–Win Denial: The Psychological Underpinnings of Zero-Sum Thinking

Supplementary Materials

Part A: Study 1 Stimuli and Results Details

	Means		Transaction Types		
Monetary Transactions	**Buyer**	**Seller**	**Win–Win**	**Win–Lose (Seller Wins)**	**Zero–Zero**
Jenny goes to Anne's grocery store. Jenny pays Anne $8 for a bottle of olive oil.	0.51	1.61	46.5%	34.9%	4.7%
Sam goes to Matt's car dealership. Sam pays Matt $33,000 for a Toyota.	–0.02	3.09	52.3%	39.5%	0.0%
Mary goes to Jack's supermarket. She pays Jack $2 for a chocolate bar.	0.65	1.19	46.5%	27.9%	7.0%
Sally goes to Tony's clothing store. She pays Tony $30 for a shirt.	0.42	2.16	53.5%	38.4%	4.7%
Eric goes to Paul's barber shop. Eric pays Paul $15 for a haircut.	1.10	1.92	60.5%	19.8%	2.3%
Jane goes to Anna's massage parlor. Jane pays Anna $60 for a 1 hour massage.	1.35	2.07	67.4%	17.4%	1.2%
Kevin goes to Betty's pet grooming center. Kevin pays Betty $45 for his dog to be groomed.	0.66	2.23	57.0%	29.1%	2.3%
Amy goes to David's plumbing company. Amy pays David $130 for her shower to be fixed.	0.81	2.36	59.3%	31.4%	2.3%
Barter Transactions	**Traders**		**Win–Win**	**Win–Lose**	**Zero–Zero**
Mark goes to his neighbor Fred. Mark trades his bottle of soy sauce for Fred's bottle of vinegar.	0.55		31.4%	15.1%	43.0%
Angela goes to her friend Clair's house. Angela trades her pencil for Clair's eraser.	0.38		32.6%	18.6%	32.6%
Maria goes to her cousin Ivy's house. Maria trades her Kate Spate crossbody bag for Ivy's Michael Kors crossbody bag.	0.63		27.9%	29.1%	34.9%
Vivian goes to her colleague Tommy's office. She trades her Burger King hamburger for Tommy's McDonald's hamburger.	0.46		29.1%	25.6%	37.2%

Table S1. Study 1 Stimuli and Results Details.

Note. Entries in the "Means" columns are the means for buyers and sellers (for the monetary transactions) and for traders (averaged across the two parties for barters). Entries in the "Transaction Types" columns are the proportion of times that participants categorized each transaction as win–win (positive scores for both parties), win–lose (for monetary transactions, positive scores for the seller and negative scores for the buyer; for barters, positive scores for one party and negative scores for the other), or zero–zero (scores of zero for both parties), calculated as in Table 1 in the main text. No other transaction type had greater than 7% prevalence for any item (see Table 1 in the main text for aggregated data).

Part B: Supplementary Studies

Study S1: Replication of Study 1 with "Benefit" Wording

One potential concern about our studies is that participants are misinterpreting the terms "better off" and "worse off." These terms are used by economists to refer to improvements and decrements in an individual's overall welfare. However, in common parlance these terms might be used at times to refer to specifically *financial* well-being, as opposed to overall well-being. If participants are misinterpreting the terms in this way, it is no surprise that they would look like mercantilists, "overvaluing" money, since they interpret the question as referring *only* to money.

To rule out this concern, we used an alternative phrasing of the key dependent measure, asking if the buyers, sellers, and traders "benefited" from the transactions. This wording in fact stacks the deck in the opposite direction, since one could (in principle) answer this question affirmatively if one thinks a buyer benefitted to some degree but not to an extent worth the cost—such a buyer still "benefited" from the transaction, just not on balance. Thus, although we anticipated that people would be more likely to give affirmative responses to this question overall, we should still see the same asymmetries among buyers, sellers, and traders seen in Study 1.

Methods. We recruited 100 participants (M_{age} = 37.6, 62% female, 55% college educated) from Mechanical Turk; 15 were excluded from analysis based on the criterion used in the other studies.

The method was identical to Study 1, except that participants were asked "To what extent do you think [*buyer / seller / trader*] benefited from this transaction?" for each transaction party, instead of the Study 1 phrasing of "How well off do you think [*buyer / seller / trader*] now is?" The scale remained anchored at –5 ("Worse than before"), 0 ("Same as before"), and 5 ("Better than before").

Results and Discussion. As shown in Figure S1, participants' welfare-change ratings were indeed more positive than in Study 1, reflecting the question asking about benefits without mentioning costs. But critically, the asymmetries among buyers, sellers, and traders remained large in magnitude.

Buyers were seen as losing from the transaction much more often than sellers [M = 1.21, SD = 1.62 vs. M = 0.35, SD = 0.77 out of 8; $t(84) = 5.15$, $p < .001$, $d = 0.68$], and this trend was also reflected in the mean welfare-change scores [M = 1.88, SD = 1.52 vs. M = 2.93, SD = 1.34; $t(84) = 7.44$, $p < .001$, $d = 0.73$]. While these asymmetries are somewhat less pronounced compared to Study 1, they are large in magnitude and reflect a general concern that buyers fail to benefit from their purchases, which we have also seen in every other study reported in this article.

Also mirroring other studies, traders were much more assigned zero-change scores [M = 1.32, SD = 2.04 out of 8] much more often than buyers [M = 0.19, SD = 0.92 out of 8; $t(84) = 5.42$, $p < .001$, $d = 0.71$] or sellers [M = 0.16, SD = 0.91 out of 8; $t(84) = 5.42$, $p < .001$, $d = 0.73$]. This also resulted in lower mean welfare-change scores for traders compared to monetary transaction parties (collapsing across buyers and sellers) [M = 1.21, SD = 1.34 vs. M = 2.41, SD = 1.28; $t(84) = 7.81$, $p < .001$, $d = 0.92$]. This confirms that traders are really seen as failing to benefit from their transactions, not merely as being no "better off" financially afterwards.

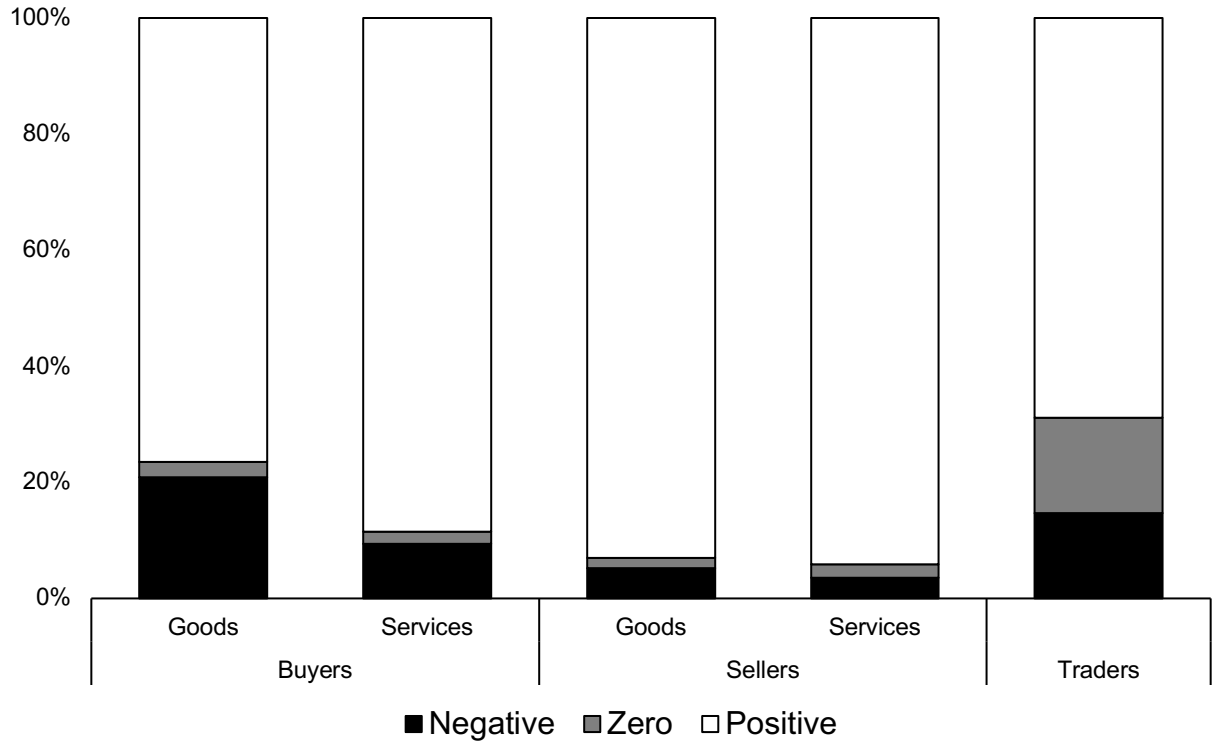

Figure S1. Proportion of transactions perceived as having negative, zero, or positive impact on buyers, sellers, and traders in Study 1.

Although participants are less likely to deny that buyers "benefit" than that they are made "better off" by their purchases, they still are much more likely to say that buyers, rather than sellers, are harmed by transactions. Likewise, although people are somewhat likelier to say that traders "benefited" from their barters rather than are made "better off," they are far likelier to say that barters fail to benefit either party compared to monetary transactions. Thus, the key asymmetries documented throughout the studies—and offered as one key piece of evidence for mercantilist intuitions—are not an artifact of the way the dependent measure is worded.

Study S2: Between-Subjects Replication of Study 2

A second piece of evidence for mercantilist thinking—beyond the asymmetries among buyers, sellers, and traders—is that reframing transactions in terms of time rather than money tends to reduce the asymmetries between buyer and seller. This follows from mercantilism, which is the systematic tendency to overvalue money relative to the goods and services it can purchase. Sellers are seen as benefitting because they gain money (overvalued relative to the goods they give up) while buyers are seen as losing because they gain goods (undervalued relative to the money they give up). When money is made less salient, it should attenuate these intuitions and reduce asymmetries between buyers and sellers, as we saw in Study 2.

Study S2 aimed to go beyond Study 2 in two respects. First, it used a between-subjects designs. Specifically, Study 2A design was identical to Study 1 except that, rather than framing all purchases in terms of money, it framed them in terms of time (as in the *time-frame* condition of Study 2). That is, comparing Study S2A to Study 1 is equivalent to testing the effect of money-frame versus time-frame (tested within-subjects in Study 2) in a between-subjects design.

Second, Study S2B tested a *money-plus-time* framing in which both the monetary and time value were specified. This helps to tease apart whether the salutary effects of time-framing are due to mentioning time, per se, or instead avoiding the mention of money. According to mercantilism it is the latter avoidance of money-talk that produces the attenuation of the asymmetries between buyer and seller. Thus, we would expect Study S2B to look more like Study 1 than like Study S2A.

Methods. We recruited 199 participants (M_{age} = 34.4, 55% female, 47% college educated) from Amazon Mechanical Turk (N = 99 for Study S2A and N = 100 for Study S2B); 55 were excluded from analysis based on the criterion used in the other studies.

Study S2A used the same design as Study 1, except every item appeared in the *time-frame* used in Study 2. That is, the buyers' side of each monetary transaction was described in terms of labor, as was one of the traders for each of the barter. For example, whereas an item in the money-frame used in Study 1 read "Sally goes to Tony's clothing store. She pays Tony $30 for a shirt," the equivalent time-frame item in Study S2A (as well as the time-frame condition of Study 2) read "Sally goes to Tony's clothing store. Sally worked for 1.5 hours to pay Tony for a shirt," with a separate paragraph describing Sally's after-tax wages at $20 per hour. One of the barter items read "Vivian goes to her colleague Tommy's office. Vivian worked for 20 minutes to pay for a Burger King hamburger, which she trades to Tommy in exchange for Tommy's McDonald's hamburger," as in the time-frame of Study 2.

Study S2B used a *money-plus-time-frame*, which differed subtly from the time-frame in Study S2A by including both the monetary *and* the time cost: "Sally goes to Tony's clothing store. She pays Tony $30 (which took Sally 1.5 hours of work to earn) for a shirt." A barter item read "Vivan goes to her colleague Tommy's office. She trades her Burger King hamburger (which took Vivian 20 minutes of work to earn) for Tommy's McDonald's hamburger." This phrasing does not explicitly mention the monetary value of the good (since the Study 1 phrasing did not do so either) but implicitly cued money by mentioning that it took a given amount of time to "earn" the good.

Results and Discussion. As shown in Figure S2, the asymmetries among buyers, sellers, and traders replicate once again, but are generally smaller in Study S2A (in the *time-frame*) compared to Study 1 (the *money-frame*) or Study S2B (the *money-plus-time-frame*). This buttresses the results of Study 2 and further implicates mercantilism as the driver of these effects by showing that it is avoiding money-talk, rather than introducing time-talk, that is driving these framing effects.

First, we note that the key asymmetries among buyers, sellers, and traders remain statistically robust in Studies S2A and S2B. Buyers were likelier to be deemed worse-off compared to sellers [$t(60) = 4.68$, $p < .001$, $d = 0.78$ and $t(82) = 6.94$, $p < .001$, $d = 1.08$ in Studies S2A and S2B, respectively], which led to higher welfare-change scores for buyers than for sellers [$t(60) = 3.60$, $p < .001$, $d = 0.49$ and $t(82) = 7.61$, $p < .001$, $d = 0.85$]. Moreover, traders were seen as likelier to experience no-benefit

compared to both buyers [$t(60) = 5.11$, $p < .001$, $d = 0.84$ and $t(82) = 9.02$, $p < .001$, $d = 1.18$] and sellers [$t(60) = 4.42$, $p < .001$, $d = 0.77$ and $t(82) = 9.22$, $p < .001$, $d = 1.27$], resulting in lower welfare-change scores on average for barters versus monetary transactions (averaging across buyers and sellers) [$t(60) = 7.03$, $p < .001$, $d = 0.60$ and $t(82) = 7.39$, $p < .001$, $d = 0.75$]. All of these results qualitatively replicate the asymmetries found in the other studies.

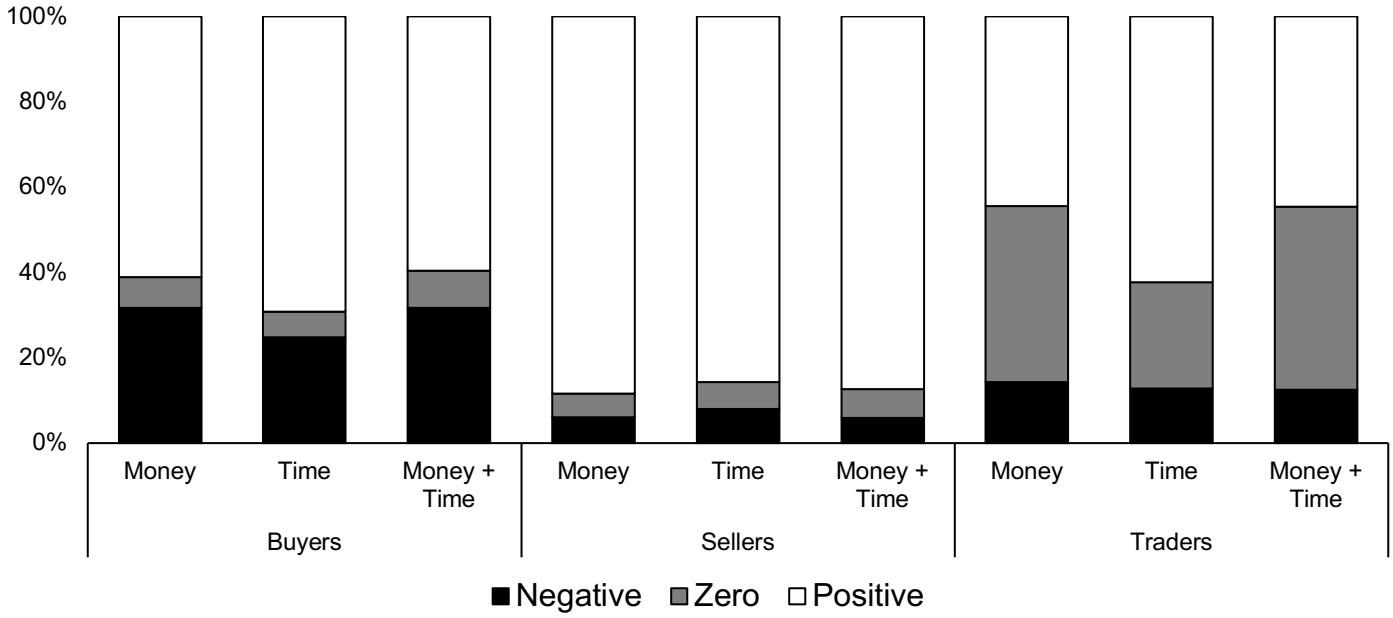

Figure S2. Proportion of transactions perceived as having negative, zero, or positive impact on buyers, sellers, and traders in the *money-frame* condition (Study 1), *time-frame* condition (Study S2A), and *money-plus-time* condition (Study S2B).

But quantitatively, there was some evidence for variation in the magnitude of these asymmetries across studies. Just as the asymmetries were smaller in the *time-frame* condition of Study 2 compared to the *money-frame* condition, the asymmetries were smaller in Study S2A than in Study 1. Some of these differences across studies reached significance and others did not, but all were directionally consistent with predictions.

Buyers were seen as directionally (but not significantly) less likely to experience negative welfare-change in Study S2A [$M = 1.98$, $SD = 2.12$ out of 8] compared to Study 1[$M = 2.53$, $SD = 2.47$ out of 8; $t(145) = 1.41$, $p = .16$, $d = 0.24$] or Study S2B [$M = 2.54$, $SD = 2.58$ out of 8; $t(142) = 1.38$, $p = .17$, $d = 0.23$]. These led to comparable (non-significant) trends on mean welfare-change scores for buyers across experiments [$M = 1.03$, $SD = 1.41$ vs. $M = 0.69$, $SD = 1.48$ vs. $M = 0.74$, $SD = 1.38$; $t(145) = 1.45$, $p = .15$, $d = 0.24$ and $t(142) = 1.28$, $p = .20$, $d = 0.22$].

Traders were seen as significantly less likely to experience no-benefit in Study S2A [$M = 2.00$, $SD = 2.37$ out of 8] compared to Study 1 [$M = 3.29$, $SD = 2.98$ out of 8; $t(145) = 2.81$, $p = .006$, $d = 0.47$] or Study S2B [$M = 3.43$, $SD = 3.01$ out of 8; $t(142) = 3.08$, $p = .002$, $d = 0.52$]. However, the differences in mean welfare-change scores for traders across experiments did not reach significance [$M = 0.71$,

SD = 1.06 vs. M = 0.51, SD = 0.93 vs. M = 0.51, SD = 0.94; $t(145) = 1.26$, $p = .21$, $d = 0.21$ and $t(142) = 1.22$, $p = .22$, $d = 0.21$].

Overall, these results—while less statistically robust than the within-subjects Study 2—help to buttress the results of that study in two ways. The money-frame versus time-frame across Studies 1 and S2A produced reduced asymmetries comparable in magnitude (if not larger) than those in Study 2, although some of these reductions did not reach statistical significance. Second, the fact that Study S2B and Study 1 produced very similar results indicates that a money-frame and money-plus-time-frame have similar psychological impact. That is a further piece of evidence for mercantilism, which predicts that cueing money (even if time is also cued) would result in perceived differences in buyer versus seller welfare.

Study S3: Between-Subjects Replication of Study 3

We argued that although mercantilism explains the asymmetries between buyers and sellers, it does not necessarily predict that they would need to be as sharp as they are across these studies. Plausibly, other factors may also be at play, and we argued that theory-of-mind or naïve realism is likely to be one such factor (Ross & Ward, 1997; Ross et al., 1977). Supporting this, Study 3 found that cueing participants to consider the buyers' reasons—even empty reasons such as because the buyer "wanted" the product—reduced win–win denial.

Study S3 had two goals. First, analogously to Study S2, it aimed to replicate Study 3 in a between-subjects design. Whereas Study 3 exposed participants to both the *empty-reason* and *no-reason* condition, Study S3A framed every item in the empty-reason condition. Thus, comparing Study S3A to Study 1 is equivalent to testing the effect of the reason versus no-reason conditions (tested within-subjects in Study 3) in a between-subjects design. Second, Study S3 tested whether different kinds of reasons have differential impact. Thus, whereas Study S3A used the same empty reasons from Study 3, Study S3B gave "content" reasons that appealed directly to the buyer's preferences.

Methods. We recruited 199 participants (M_{age} = 38.8; 67% female; 55% college educated) from Amazon Mechanical Turk (N = 99 for Study S3A and N = 100 for Study S3B); 26 were excluded from analysis based on the criterion used in the other studies.

Study S3A used the same design as Study 1, except every item appeared in the *empty-reason* phrasing used in Study 3. That is, each transaction included an empty explanation of the buyer's or traders' choice, such as "Sally made the purchase because she wanted the shirt." Study S3B was identical, except that a *content-reason* was given instead, which appealed to the buyer's or traders' preferences, such as "Sally purchased the shirt because Taylor Swift once wore this kind of shirt at her concert, and Sally loves Taylor Swift very much" or "Eric got the haircut because he finds the environment in Paul's barber shop to be pleasant and enjoyable." Similarly, for one of the barters, "Mark traded because he needed vinegar for a recipe, and Fred traded because he happened to run out of soy sauce."

Results and Discussion. As shown in Figure S3, we see yet again the asymmetries among buyers, sellers, and traders predicted by mercantilism. But these are less statistically robust than in other

experiments and do not always reach significance—this is the only study in the main or supplementary text that does not produce highly robust asymmetries between traders and monetary transactors. This is because win–win denial was overall much lower in the empty-reason and content-reason versions used in Study S3, compared to the no-reason version used in Study 1. Indeed, these differences among conditions were much larger in magnitude than in the within-subjects Study 3. These results point to theory-of-mind errors as a partial driver of win–win denial and to perspective-taking as a useful corrective.

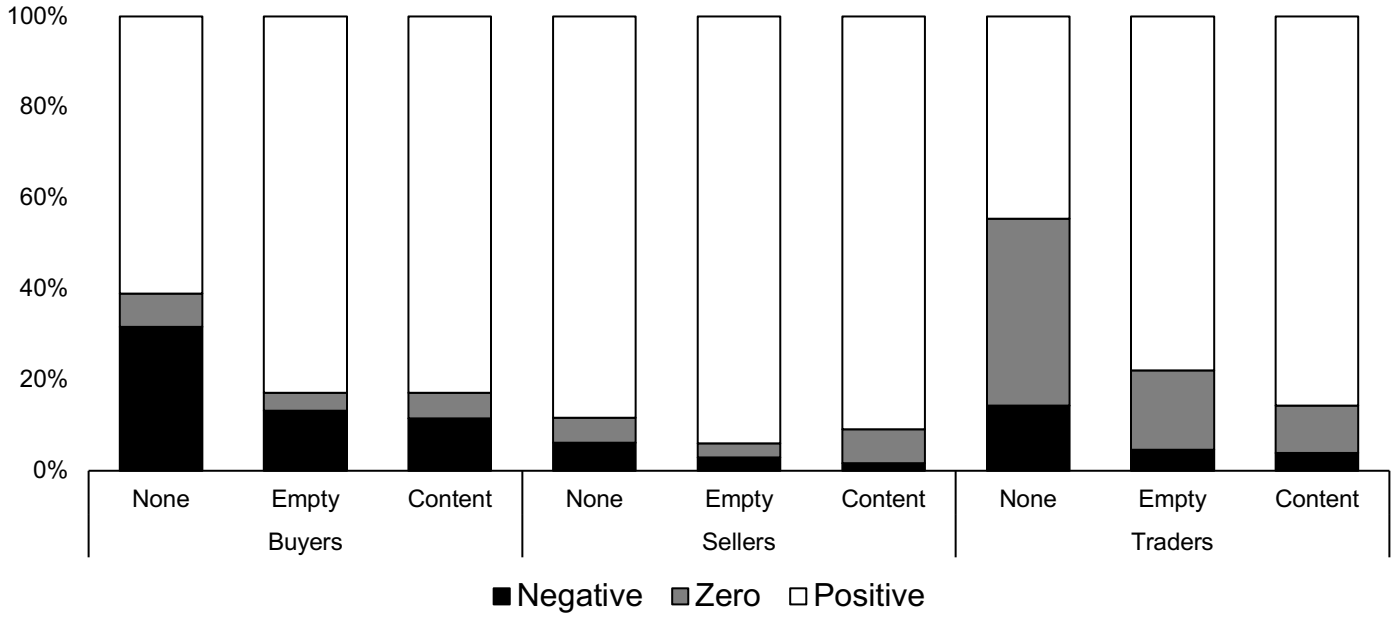

Figure S3. Proportion of transactions perceived as having negative, zero, or positive impact on buyers, sellers, and traders in the *no-reason* condition (Study 1), *empty-reason* condition (Study S3A), and *content-reason* condition (Study S3B).

First, we examine again the asymmetries among buyers, sellers, and traders. Buyers remained likelier to be seen as made worse-off compared to sellers in both Study S3A [$t(86) = 4.08$, $p < .001$, $d = 0.56$] and Study S3B [$t(85) = 4.95$, $p < .001$, $d = 0.73$], with these differences also reflected in the mean welfare-change scores in both studies [$t(86) = 4.95$, $p < .001$, $d = 0.38$ and $t(85) = 4.06$, $p < .001$, $d = 0.37$]. In both studies, traders were likelier to be given no-change scores compared to both buyers [$t(86) = 3.79$, $p < .001$, $d = 0.57$ and $t(85) = 2.01$, $p = .048$, $d = 0.25$] and sellers (though not significantly for Study S3B) [$t(86) = 4.52$, $p < .001$, $d = 0.61$ and $t(85) = 1.24$, $p = .22$, $d = 0.14$]. Unlike the difference between buyers and sellers, the mean welfare-change scores did not significantly differ between barters and monetary transactions (averaging across buyers and sellers) in either experiment [$t(86) = 1.08$, $p = .28$, $d = 0.08$ and $t(85) = 1.24$, $p = .22$, $d = 0.10$]. As we will see, this reflects the much lower rate of win–win denial seen in Study S3.

Next, we directly compare the rates of win–win denial across studies. Buyers were seen as losing from their transactions much less frequently, compared with the no-reason framing in Study 1 [$M = 2.53$, $SD = 2.47$ out of 8], when empty-reasons were given in Study S3A [$M = 1.06$, $SD = 1.98$ out of

8; $t(171) = 4.34$, $p < .001$, $d = 0.66$] or content-reasons were given in Study S3B [$M = 0.92$, $SD = 1.47$ out of 8; $t(170) = 5.21$, $p < .001$, $d = 0.79$]; Studies S3A and S3B did not differ from one another [$t(171) = 0.52$, $p = .60$, $d = 0.08$]. These trends were also reflected in the mean welfare-change scores for buyers, which were significantly more positive in Study S3A [$M = 2.18$, $SD = 1.79$; $t(171) = 5.98$, $p < .001$, $d = 0.91$] and Study S3B [$M = 2.16$, $SD = 1.35$; $t(170) = 6.82$, $p < .001$, $d = 1.04$] compared to Study 1 [$M = 0.69$, $SD = 1.48$].

Similar effects were seen for traders. The perception that traders experience no change from their barters was significantly lower in both Study S3A [$M = 1.40$, $SD = 2.58$ out of 8; $t(171) = 4.46$, $p < .001$, $d = 0.68$] and Study S3B [$M = 0.84$, $SD = 1.84$ out of 8; $t(170) = 6.50$, $p < .001$, $d = 0.99$], relative to Study 1 [$M = 3.29$, $SD = 2.98$]; this effect was slightly larger in Study S3B (with content reasons), reflected in a marginally significant difference between Studies S3A and S3B [$t(171) = 1.66$, $p = .099$]. These effects translated into much higher welfare-change scores for barters in Study S3A [$M = 2.34$, $SD = 1.92$; $t(171) = 7.98$, $p < .001$, $d = 1.21$] and Study S3B [$M = 2.54$, $SD = 1.45$; $t(170) = 10.96$, $p < .001$, $d = 1.67$] compared to Study 1 [$M = 0.51$, $SD = 0.93$].

Overall, the perspective-taking manipulations led to much lower rates of win–win denial and zero-sum thinking. Whereas most participants in Study 1 (88%) believed that at least one of the transactions was non-positive-sum, this proportion was far lower in Studies S3A (39%) and S3B (41%). Likewise, whereas nearly all participants in Study 1 (94%) believed that at least one individual in one of the transactions failed to benefit, this proportion was more modest in Studies S3A (60%) and S3B (65%). Thus, although perspective-taking interventions certainly did not *eliminate* zero-sum thinking, it dramatically lowered its incidence.

Overall, these results attest to the significance of theory-of-mind in win–win denial—providing reasons for buyers' choices, even if totally devoid of content, cues people to realize that people rarely voluntarily trade at a loss. Nonetheless, theory-of-mind cannot be the whole story, since it does not predict the asymmetries that we are consistently seeing among buyers, sellers, and traders. Theory-of-mind errors and intuitive mercantilism work together to produce our flawed intuitions about exchanges, and both appear to contribute to zero-sum thinking.

Study S4: Cueing Preferences

Study S4 further tested the theory-of-mind account by examining whether increasing the saliency of the transactors' preferences increases the perception that the exchanges are win–win. That is, it is likely that, if asked explicitly, people would recognize that buyers preferred the good they purchased over the money they spent, while sellers prefer the money over the good. Indeed, neoclassical economics goes so far as to *define* preferences as revealed by one's choices (Mas-Colell et al., 1995). Even if people do not hold this strong assumption of revealed preferences, they probably recognize that choices are typically aligned with preferences. However, from the fact that preferences were satisfied it follows logically that both parties to each exchange benefitted. Hence, cueing preferences as an intermediate step may help to circumvent theory-of-mind limits.

A second goal of Study S4 was to further rule out the concern that participants interpret the dependent variable in exclusively monetary terms. This was addressed partly in Study S1, which used

the phrase "benefitted" rather than "better off," but we put this concern to rest in Study S4 by providing participants with detailed instructions about how we define "better off" and requiring participants to correctly answer a series of check questions to ensure that they can correctly apply this definition.

Methods. We recruited 100 participants (M_{age} = 43.5, 49% female, 58% college educated) from Mechanical Turk; 19 were excluded from analysis based on the criterion used in the other studies (failing at least 33% of the check questions at the end of the study) or because they answered one or more of the instruction comprehension check questions incorrectly. A further 12 participants were excluded who made an error on any of the preference questions (see below). Hence, all participants included in the final sample acknowledged that all transactions satisfied the preferences of both parties. As shown below, however, excluding participants based on their answers to the preference questions does not affect the results.

The procedure was similar to Study 1, with two changes. First, participants were provided detailed definitions of the terms "better off" and "worse off":

> For each transaction, you will be asked whether each participant is "better off," "worse off," or "the same," relative to how they were before the transaction. When we use these terms, we are asking whether, all things considered, each person has benefited – in the language of economics, whether each person has increased "utility" after the transaction compared to beforehand. We are using a very wide definition of "benefit" or "utility" that includes any monetary gains or losses and any gains or losses in convenience, health, pleasure, or any other reason a person would carry out a transaction.

Then participants completed four sample items that were framed around chance events rather than transactions, so as not to introduce demand characteristics for the main task. These events were "George finds a $10 bill on the ground but loses a $1 bill," "Rachel finds a $5 bill on the ground but loses a $100 bill," "Ralph finds a $100 bill on the ground but lost a movie ticket that he would have been willing to pay $10 for," and "Judy finds a $5 bill on the ground but lost a concert ticket that she would have been willing to pay $50 for." For each event, participants responded as a forced-choice whether each person was made "better off," "the same," or "worse off." Crucially, two of these items required participants to compare monetary and non-monetary utility, and in one case the correct answer was that the person was made worse-off despite gaining currency. Participants incorrectly answering any of these questions were excluded from data analysis, implementing a maximally strict standard for comprehension of the dependent variable.

Second, for half of the items (in the *preference-cue* condition), participants were first asked to state (as a forced-choice) the preferences of each party to each transaction. For example, for the item asking about Sally's trip to Tony's clothing store to buy a shirt for $30, participants were asked "Which do you think Sally preferred?" and "Which do you think Tony preferred?" as a forced-choice ("the shirt" or "$30"). Likewise, for the item asking about Tommy and Vivian's swap of their Burger King and McDonald's burgers, participants were asked which Tommy and Vivian preferred ("the Burger King hamburger" or "the McDonald's hamburger"). Participants were highly accurate in answering these questions. Among participants who were not excluded due to errors on the instruction or memory check questions, the average number of errors on these questions was 0.35 out of 8, with all but 12

participants answering all questions correctly. For the other half of the items (in the *no-cue* condition), these preference questions were omitted. The 4 items in each condition were blocked and the order of the two blocks was counterbalanced.

Because the preference questions are difficult to phrase unambiguously for services (e.g., which did the dog groomer prefer – $45 or the time and labor associated with grooming the dog), we omitted these items to ensure that the manipulation was clear. Hence, participants in Study S4 saw the 4 monetary purchases of goods and 4 barters used in the other studies, with half of the monetary purchases and half of the barters appearing each in the preference-cue and no-cue conditions, counterbalanced across items.

Results. Overall, the no-cue condition replicates the mercantilist pattern seen in the other studies. This pattern is attenuated in the preference-cue condition, just as cueing buyers' reasons for transactions in Studies 3 and S3 reduced win–win denial.

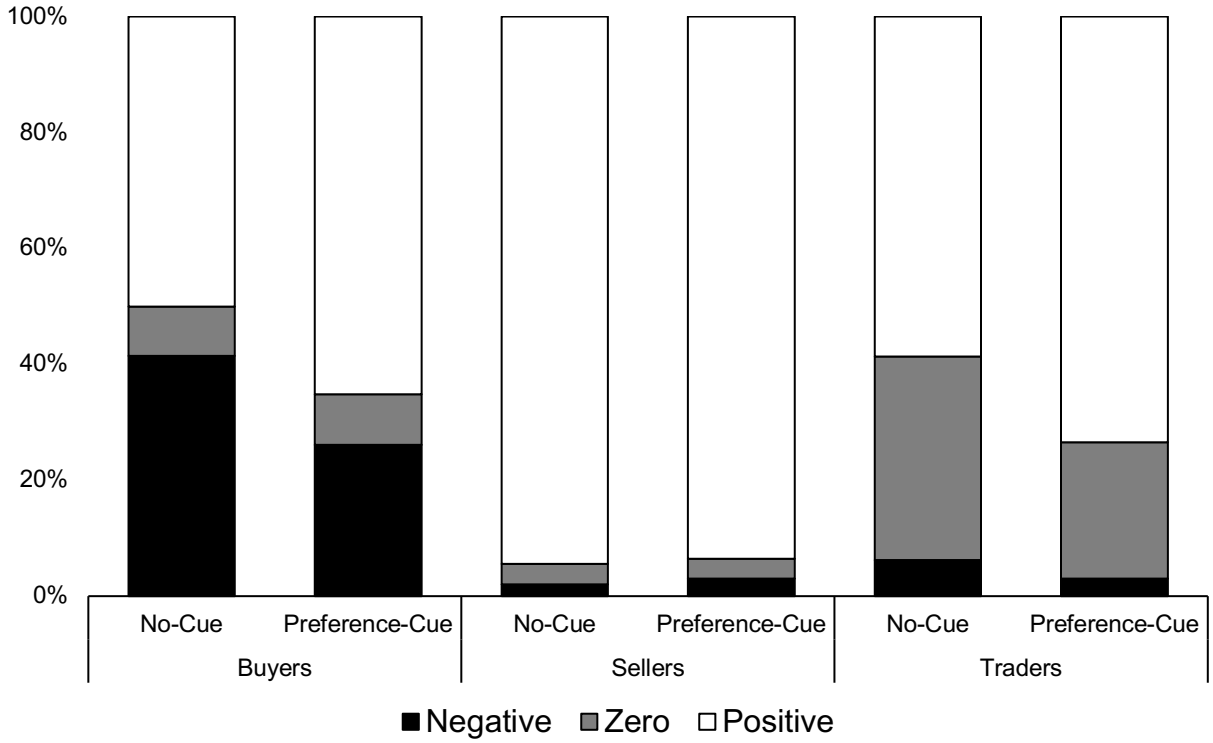

Figure S4. Proportion of transactions perceived as having negative, zero, or positive impact on buyers, sellers, and traders across no-cue and preference-cue conditions of Study S4.

We first examine the no-cue condition, which is conceptually similar to Study 1 in the main text, except that participants were inculcated in the meaning of "better off" and "worse off." The results were very similar to those of Study 1. As shown in Figure S4, in the no-cue condition, participants much more frequently claimed that buyers rather than sellers were harmed by transactions [$M = 0.83$, $SD = 0.86$ vs. $M = 0.04$, $SD = 0.21$ out of 2; $t(68) = 7.45$, $p < .001$, $d = 1.26$]. This led to lower welfare scores for buyers than for sellers [$M = 0.04$, $SD = 1.80$ vs. $M = 1.85$, $SD = 1.42$; $t(68) = 6.63$, $p < .001$, $d = 1.12$].

To adjust for the fact that Study S4 included 4 observations of trader welfare in each condition (2 observations for each of 2 items) but only 2 observations of buyer and seller welfare in each condition (1 observation each of buyer and seller welfare for each of 2 items), we divide the counts of positive, zero, and negative welfare change for traders to create comparable measurements. Participants far more frequently claimed that traders [$M = 0.70$, $SD = 0.84$ out of 2 after adjustment] experience zero change in welfare compared to buyers [$M = 0.17$, $SD = 0.48$ out of 2; $t(68) = 5.00$, $p < .001$, $d = 0.77$] or sellers [$M = 0.07$, $SD = 0.31$ out of 2; $t(68) = 5.70$, $p < .001$, $d = 1.00$]. This manifested in marginally lower welfare-change scores for barters than for monetary transactions (averaging across buyers and sellers) [$M = 0.69$, $SD = 1.27$ vs. $M = 0.94$, $SD = 1.16$; $t(68) = 1.73$, $p = .088$, $d = 0.20$]. Thus, the mercantilist intuitions on evidence in the other studies are replicated in Study S4, even with detailed definitions of "welfare" and instruction check questions to ensure comprehension of this definition.

Next, we compare the preference-cue to the no-cue condition. If the preference-cue assists participants in taking the perspective of the buyers and traders, then we would expect buyers and traders to be seen as more likely to gain, and hence their welfare-change scores to be higher. We would not necessarily expect comparable effects for sellers, for the reasons given in connection with Study 3, which revealed only weak tendencies for mentalizing to increase the perceived welfare of sellers.

As shown in Figure S4, these predictions were borne out. In the preference-cue condition, buyers were less likely to be seen as losing [$M = 0.52$, $SD = 0.78$ vs. $M = 0.83$, $SD = 0.86$ out of 2; $t(68) = 3.65$, $p < .001$, $d = 0.37$] and traders less likely to be seen as experiencing zero change [$M = 0.94$, $SD = 1.49$ vs. $M = 1.41$, $SD = 1.67$ out of 2; $t(68) = 3.33$, $p = .001$, $d = 0.29$], but sellers equally likely to be seen as gaining [$M = 1.87$, $SD = 0.45$ vs. $M = 1.88$, $SD = 0.37$ out of 2; $t(68) = 0.24$, $p = .81$, $d = 0.04$], compared to the no-cue condition. This manifested in more positive welfare-change scores for buyers [$M = 0.59$, $SD = 1.65$ vs. $M = 0.04$, $SD = 1.80$; $t(68) = 3.01$, $p = .004$, $d = 0.32$] and traders [$M = 1.20$, $SD = 1.63$ vs. $M = 0.69$, $SD = 1.27$; $t(68) = 4.03$, $p < .001$, $d = 0.34$], but not sellers [$M = 1.79$, $SD = 1.43$ vs. $M = 1.85$, $SD = 1.42$; $t(68) = –0.30$, $p = .77$, $d = –0.04$].

Plausibly, one might be concerned that excluding participants who made any errors on the preference questions could lead to a differential attrition problem that would bias the results in favor of more positive scores in the preference-cue condition. That is, since the preference questions were only asked in the preference-cue condition, one could argue that we selectively eliminated participants who had particularly negative views of the transactions in that condition. Hence, we repeated all analyses adding back in the participants who were eliminated due to their answers to these questions ($N = 12$). The results were very similar. In the preference-cue condition, buyers were less likely to be seen as losing [$t(80) = 3.79$, $p < .001$, $d = 0.36$] and traders less likely to be seen as experiencing zero change [$t(80) = 3.29$, $p = .001$, $d = 0.25$], but sellers no less likely to be seen as gaining [$t(80) = 0.00$, $p = 1.00$, $d = 0.00$]. Once again this resulted in more positive welfare-change scores for buyers [$t(80) = 3.43$, $p < .001$, $d = 0.34$] and traders [$t(80) = 4.04$, $p < .001$, $d = 0.31$] but not sellers [$t(80) = 0.44$, $p = .66$, $d = 0.06$].

Potentially we might see carry-over effects from one part of the experiment to the next. That is, participants may be more prone to seeing transactions as win–win even in the no-cue condition, if this condition had been preceded by the preference-cue condition. However, the extent to which buyers were seen as losing, sellers as winning, and traders as experiencing zero-change did not differ across

the two orders (ts < 1.2, ps > .27), nor did the mean welfare-change scores for buyers, sellers, or traders (ts < 1.7, ps > .10).

Discussion. Study S4 supports three conclusions. First, participants rarely claimed that the buyers or traders did not prefer the item that they in fact received, suggesting that our minds make a strong intuitive link between preferences and choices. This cuts somewhat against the interpretation that participants believe the buyers or traders are irrational or deceived by the other party. Second, despite detailed instructions and comprehension checks about the definitions of "better off" and "worse off," participants in the no-cue condition showed similar mercantilist patterns as in the other studies. Third, these mercantilist intuitions were weaker in the preference-cue condition, supporting the idea that theory-of-mind limitations contribute to win–win denial.

Study S5: Subjective Beliefs and First-Person Perspective

Given the large effects of cueing reasons and preferences in Studies 3, S3, and S4, we considered two other potential manipulations that might influence the extent of win–win denial.

Study S5A asked participants to judge the buyers' and sellers' *beliefs* about their relative welfare. By asking participants to directly judge whether, for example, Sally *believes* she is bettered by her chocolate bar purchase, they may be encouraged to spontaneously consider Mary's reasons for making the purchase, which would lead to a pattern more like Studies 3 and S3 (where reasons were provided) than Study 1 (where they were not). This would also suggest that people believe that others are *mistaken* when they engage in exchanges—they may believe that they are better off, but they in fact are not. This is the prediction made by the theory-of-mind account, according to which people typically substitute their preferences for that of the buyer. Asking people to explicitly think from the buyer's perspective by using a subjective question wording should, on this account, have a salutary effect on win–win denial.

Study S5B placed the participant in the position of the buyer by adopting a first-person perspective. On the one hand, people are known to exaggerate their decision-making abilities (e.g., Johnson & Rips, 2014, 2015; Pronin & Kugler, 2010) and might therefore be likelier to think themselves as gaining from their purchases. But on the other hand, people actually know their own preferences, and in many cases the transactions really may not be worth it. You aren't better-off spending $2 for a chocolate bar if you don't like chocolate! If theory-of-mind errors are at work, these first-person judgments might look very similar to third-person judgments, since one might impute one's own anti-chocolate preferences onto the third-person buyer. Thus, this account predicts that Study S5B would produce similar results to Study 1.

Methods. We recruited 199 participants (M_{age} = 37.9; 61% female; 43% college educated) from Amazon Mechanical Turk (N = 100 for Study S5A and N = 99 for Study S5B); 31 were excluded from analysis based on the criterion used in the other studies.

For Study S5A, the procedure was identical to Study 1, except that rather than rating the buyers' and sellers' welfare, participants rated the buyers' and sellers' *perceived* welfare. For example, the

dependent measures for Sally and her clothing purchase from Tony were "How well off do you think Sally believes she now is?" and "How well off do you think Tony believes he now is?" The instructions were modified to reflect this change to a subjective wording ("For each transaction, you will be asked whether each participant believes that he or she is better off, worse of, or the same, relative to how they were before the transaction").

For Study S5B, the procedure was identical to Study 1, except that the buyers in the transactions were first-person ("you") rather than third-person (e.g., "Sally" or "Eric"). One item had to be lightly rephrased to make it gender-neutral, but otherwise all items were unchanged. For example, one item read "You go to Tony's clothing store. You pay Tony $30 for a shirt," and the dependent measures were "How well off do you think you now are?" and "How well off do you think Tony now is?" on the same scale as previous experiments. The instructions were modified to reflect these changes ("…you will read about some transactions, each involving you and another person").

Results and Discussion. As shown in Figure S5, win–win denial was somewhat less common in Study S5A, when explicitly taking the buyers', sellers', and traders' perspectives, compared to other studies where one is not cued to take the transactors' perspectives. However, the first-person framing in Study S5B led to similar rates of win–win denial as in Study 1, further implicating theory-of-mind. In both studies, asymmetries among buyers, sellers, and traders remained robust.

First, we replicated the asymmetries among buyers, sellers, and traders that we have seen in every other study. Buyers were seen as likelier to lose from transactions compared to sellers in both Study S5A [$t(81) = 5.41$, $p < .001$, $d = 0.76$] and Study S5B [$t(85) = 7.96$, $p < .001$, $d = 1.15$], and these trends were also reflected in mean welfare-change scores [$t(81) = 6.89$, $p < .001$, $d = 0.65$ and $t(85) = 8.89$, $p < .001$, $d = 0.99$]. Likewise, traders were seen as likelier to experience no change in welfare compared to both buyers [$t(81) = 5.15$, $p < .001$, $d = 0.69$ and $t(85) = 8.08$, $p < .001$, $d = 1.12$] and sellers [$t(81) = 5.11$, $p < .001$, $d = 0.66$ and $t(85) = 8.53$, $p < .001$, $d = 1.22$]. This was also reflected in the mean welfare-change scores, which were lower for barters than for monetary transactions (averaging across buyers and sellers) [$t(81) = 3.59$, $p < .001$, $d = 0.40$ and $t(85) = 7.95$, $p < .001$, $d = 0.77$].

The main issue of interest here is the comparison between these studies and Study 1, which used the third-person framing. Win–win denial was somewhat less common in Study S5A than in Study 1, reflecting the effect of the subjective wording. Buyers were seen as less likely to lose in Study S5A [$M = 1.67$, $SD = 2.22$ out of 8] compared to Study 1 [$M = 2.53$, $SD = 2.47$ out of 8; $t(166) = 2.38$, $p = .019$, $d = 0.37$], which was also reflected in the mean welfare-change scores [$M = 1.30$, $SD = 1.53$ vs. $M = 0.69$, $SD = 1.48$; $t(166) = 2.67$, $p = .008$, $d = 0.41$]. Analogously, traders were seen as less likely to experience no change in welfare in Study S5A [$M = 1.85$, $SD = 2.64$ out of 8] than in Study 1 [$M = 3.29$, $SD = 2.98$ out of 8; $t(166) = 3.30$, $p = .001$, $d = 0.51$], also reflected in the mean welfare-change scores [$M = 1.25$, $SD = 1.27$ vs. $M = 0.51$, $SD = 0.93$; $t(166) = 4.34$, $p < .001$, $d = 0.67$]. These results support the theory-of-mind account, in that subjectively wording the questions to emphasize the buyers' and traders' perspectives reduced win–win denial.

However, win–win denial differed little between Study S5B and Study 1, which varied only in using first-person versus third-person wording. Buyers were seen as equally likely to lose from their transactions across the two studies [$t(170) = 0.59$, $p = .56$, $d = 0.09$], resulting in similar mean welfare-

change scores [$t(170) = 0.94$, $p = .35$, $d = 0.14$]. Traders were seen as somewhat less likely to experience no change in welfare in Study S5B than in Study 1 [$M = 2.47$, $SD = 2.30$ vs. $M = 3.29$, $SD = 2.98$ out of 8; $t(170) = 2.04$, $p = .043$, $d = 0.31$], but this did not translate into a significant difference in mean welfare-change scores [$t(170) = 1.17$, $p = .24$, $d = 0.18$]. Overall, Figure S5 shows that the distributions are fairly similar between Studies 1 and S5B, which supports the theory-of-mind account by suggesting that participants are substituting a question about buyers' preferences with a question about their own.

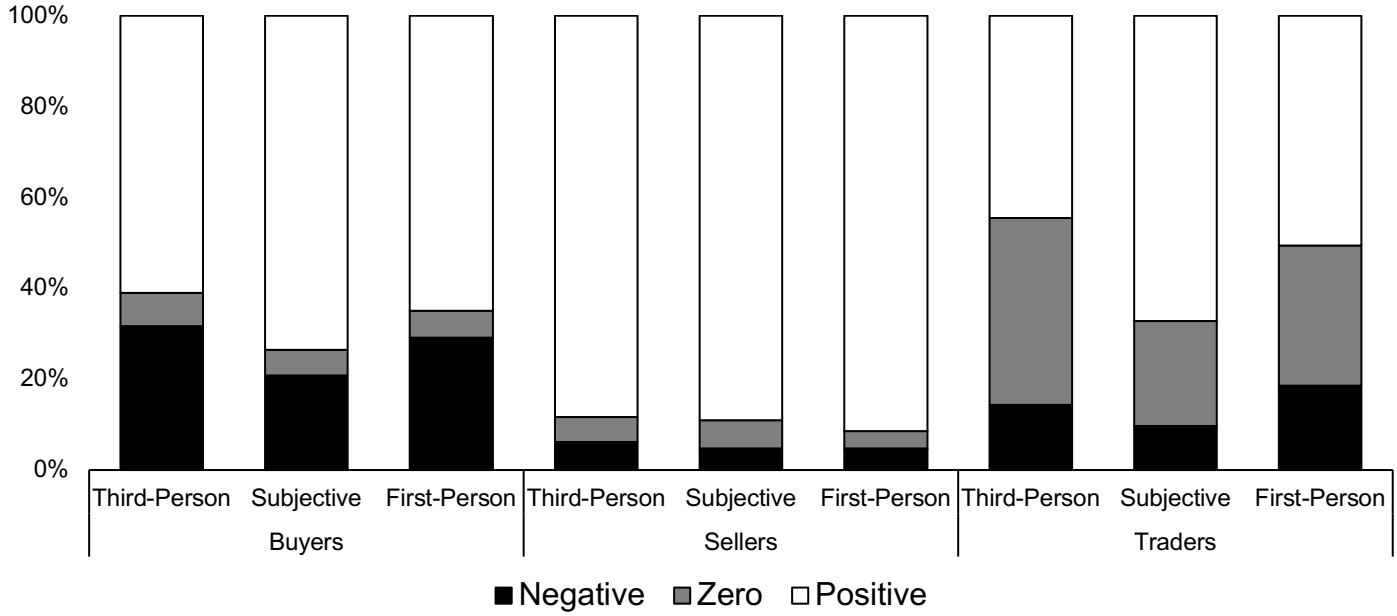

Figure S5. Proportion of transactions perceived as having negative, zero, or positive impact on buyers, sellers, and traders when transactions were framed in the *third-person* (Study 1), in terms of the transactors' *subjective* preferences (Study S5A), and in the *first-person* (Study S5B).

Overall, these results provide further support for theory-of-mind errors as a partial explanation of win–win denial. Study S5A tests another way of cueing perspective-taking, finding similar results (albeit smaller in magnitude) to Studies 3 and S3, which provided a more direct route into the buyers' mental states by giving their reasons. Study S5B, on the other hand, framed transactions in the first person and found that people answer these first-person questions in a very similar way to third-person questions, suggesting that they substitute their own preferences as they evaluate whether others gain or lose from their transactions.

As one of the reasons why markets are possible is that we all place different subjective valuations on things, this finding represents a fundamental misconception about transactions and helps to explain the incidence of zero-sum thinking. If two parties have identical preferences and make a trade, one of them is winning and the other is losing, whereas if they have different preferences (e.g., because one of them can acquire the good at a lower price or because she gains less utility than the other from the product), positive-sum transactions occur. Zero-sum thinking appears to be in part a failure of perspective-taking.

Part C: Individual Differences

Each experiment measured participants' educational background and political views. This allowed us to test hypotheses about what traits are associated with zero-sum thinking. For instance, one might imagine that economics training would attenuate zero-sum thinking, and one could plausibly argue for either liberals *or* conservatives espousing zero-sum attitudes more often (see Davidai & Ongis, 2019).

To test the impact of these variables, we built a hierarchical linear model, using the data from all experiments in the main text and Supplementary Materials (total $N = 1188$), nesting subjects within experiments and allowing a random intercept (but not random slopes) within each experiment. The dependent measure in this model was the frequency with which all parties (buyers, sellers, and traders) were perceived to be non-beneficiaries (i.e., the proportion of individuals rated as having zero or negative gains from trade, multiplied by 100). Dummy-coded variables were entered for experiment comparing these experiments to Study 1. The other variables in the model included educational variables—educational level (on a 7-point scale; *Mdn* = some college), major (dummy-coded with 1 = economics or business; 17% business/econ majors), economics knowledge (on a 0–10 scale; $M = 4.53$, $SD = 2.49$), and number of economics courses ($M = 1.40$, $SD = 1.84$); political views—social conservatism (on a 0–10 scale; $M = 3.91$, $SD = 2.95$) and economic conservatism (on a 0–10 scale; $M = 4.65$, $SD = 2.93$); and basic demographics—gender (dummy-coded with 1 = female; 59% female), age ($M = 37.9$, $SD = 12.9$), and income (square-root transformed; *Mdn* = $35,000). We also measured political party (on a 0–10 scale with higher numbers indicating stronger Republican identification; $M = 4.03$, $SD = 3.00$), but did not enter this variable into the model due to its multicollinearity with the other political variables. However, either adding this variable to the model or replacing the other two political variables with political party do not alter the results.

The model coefficients are reported in Table S2. Model 1 tests the impact of economics knowledge and education. Although formal education did not have an impact on zero-sum thinking (either overall education level or number of economics courses), self-reported economics knowledge had an appreciable effect on zero-sum thinking. For every one point (on a 0–10 scale) of self-reported economics knowledge, participants indicated about one percentage point fewer individuals had failed to benefit from their transactions. However, this effect was somewhat cancelled out by the impact of being a business or economics major, who exhibited significantly *higher* degrees of zero-sum thinking (most of these were business, rather than economics majors, within our sample). This could be (regrettably) an effect of the business curriculum at management schools, but it also may be a self-selection bias. In any event, business majors also reported higher economics knowledge (in a separate model predicting economics knowledge), helping to compensate for this tendency.

Model 2 adds political orientation—separately for economic and social conservatism—to the model. Neither variable was a significant predictor. On the one hand, this may be surprising given current political discourse, which (at the time of writing) seems to include more zero-sum language from conservative politicians, at least within the United States. But given the global resurgence of both left-wing and right-wing economic populism, it may be less surprising that political orientation is not a good predictor of a zero-sum mentality. One limitation of our measure of economic conservatism is that this may be quite conceptually distinct from support of free markets, as Caplan (2007) finds

that both political conservatives and liberals are much more skeptical about the effectiveness of markets relative to professional economists.

	Coefficient (SE)		
	Model 1	**Model 2**	**Model 3**
Education	–0.26 (0.41)	–0.22 (0.42)	–0.39 (0.43)
Business/Econ Major	3.82 (1.83)*	3.77 (1.84)*	3.13 (1.87)°
Econ Knowledge	–0.62 (0.28)*	–0.63 (0.29)*	–0.58 (0.30)°
Econ Courses	–0.43 (0.43)	–0.41 (0.43)	–0.35 (0.44)
Social Conservatism		0.22 (0.35)	0.18 (0.35)
Economic Conservativism		–0.12 (0.35)	–0.14 (0.36)
Female			1.49 (1.32)
Age			0.044 (0.05)
Income (square-root)			0.008 (0.01)

*** $p < .001$ ** $p < .01$ * $p < .05$ ° $p < .10$

Table S2. Predictors of zero-sum thinking pooling data across studies.
Note. Coefficients (SEs) predicting the proportion (multiplied by 100) of parties regarded as non-beneficiaries.

Another possibility, however, is that the political dynamics are somewhat more complex than these models suggest. Economic conservatism, social conservatism, and political conservatism were all highly correlated in our sample [rs > .79, ps < .001]. But they do appear to be conceptually distinct. For example, economic conservatism was *positively* associated with both self-reported economics knowledge [b = 0.09, p = .019] and formal economics coursework [b = 0.15, p < .001], while social conservatism had no association with economics knowledge [b = –0.00, p = .93] and a *negative* relationship with economics coursework [b = –0.08, p = .011]. Given the link between economics knowledge and lower levels of win–win denial, it could be that social conservatism and economic conservatism have opposite effects on zero-sum thinking, which are difficult to detect because they are so strongly correlated in our politically partisan environment.

Finally, Model 3 added demographic variables (gender, age, income) to the model, none of which had any additional predictive value. This is important to the extent that we wish to generalize the current results beyond the sampled population of Mechanical Turk workers, as one might argue that this population is relatively low in socioeconomic status and thus more prone to perceiving themselves as made worse-off by economic forces. But we see little evidence for this story, given no link between income and win–win denial within our sample.

Despite some significant predictors across these models, what is perhaps most striking is how *little* of the variability is explained by individual difference variables, even in cases where the factors had a statistically reliable effect. At least among our American participants, zero-sum thinking appears to generalize across people with very different political beliefs, educational backgrounds, and income levels. Apparently the zero-sum misconception is deeply rooted in human beings, and even extensive economics education does less than one would hope to root it out. The various manipulations explored in the current studies may be a more effective wedge to influence beliefs and behaviors.